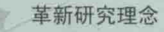


格致方法·定量研究系列 吴晓刚 主编

多元回归中的交互作用

[美] 詹姆斯·杰卡德（James Jaccard）
罗伯特·图里西（Robert Turrisi） 著

蒋勤 译



革新研究理念

丰富研究工具

最权威、最前沿的定量研究方法指南

04

格致出版社 上海人民出版社

格致方法 · 定量研究系列　吴晓刚　主编

多元回归中的交互作用

[美] 詹姆斯 · 杰卡德(James Jaccard)
罗伯特 · 图里西(Robert Turrisi) 著
蒋　勤 译

SAGE Publications ,Inc.

格 致 出 版 社　上海人民出版社

图书在版编目(CIP)数据

多元回归中的交互作用/(美)杰卡德(Jaccard，J.)，(美)图里西(Turrisi，R.)著；蒋勤译.—上海：格致出版社：上海人民出版社，2015
(格致方法·定量研究系列)
ISBN 978-7-5432-2597-8

Ⅰ.①多… Ⅱ.①杰… ②图… ③蒋… Ⅲ.①多元回归分析-研究 Ⅳ.①O212.1

中国版本图书馆 CIP 数据核字(2015)第 295286 号

责任编辑 顾 悦
美术编辑 路 静

格致方法·定量研究系列

多元回归中的交互作用

[美]詹姆斯·杰卡德 罗伯特·图里西 著
蒋勤 译

出 版	世纪出版股份有限公司 格致出版社 世纪出版集团 上海人民出版社 (200001 上海福建中路193号 www.ewen.co)	**印 刷**	浙江临安曙光印务有限公司
	编辑部热线 021-63914988 市场部热线 021-63914081 www.hibooks.cn	**开 本**	920×1168 1/32
		印 张	4.75
		字 数	90,000
		版 次	2016年1月第1版
发 行	上海世纪出版股份有限公司发行中心	**印 次**	2016年1月第1次印刷



ISBN 978-7-5432-2597-8/C·140 定价：25.00元

出版说明

由香港科技大学社会科学部吴晓刚教授主编的“格致方法·定量研究系列”丛书，精选了世界著名的 SAGE 出版社定量社会科学研究丛书，翻译成中文，起初集结成八册，于 2011 年出版。这套丛书自出版以来，受到广大读者特别是年轻一代社会科学工作者的热烈欢迎。为了给广大读者提供更多的方便和选择，该丛书经过修订和校正，于 2012 年以单行本的形式再次出版发行，共 37 本。我们衷心感谢广大读者的支持和建议。

随着与 SAGE 出版社合作的进一步深化，我们又从丛书中精选了三十多个品种，译成中文，以飨读者。丛书新增品种涵盖了更多的定量研究方法。我们希望本丛书单行本的继续出版能为推动国内社会科学定量研究的教学和研究作出一点贡献。

总 序

2003年，我赴港工作，在香港科技大学社会科学部教授研究生的两门核心定量方法课程。香港科技大学社会科学部自创建以来，非常重视社会科学研究方法论的训练。我开设的第一门课“社会科学里的统计学”(Statistics for Social Science)为所有研究型硕士生和博士生的必修课，而第二门课“社会科学中的定量分析”为博士生的必修课(事实上，大部分硕士生在修完第一门课后都会继续选修第二门课)。我在讲授这两门课的时候，根据社会科学研究生的数理基础比较薄弱的特点，尽量避免复杂的数学公式推导，而用具体的例子，结合语言和图形，帮助学生理解统计的基本概念和模型。课程的重点放在如何应用定量分析模型研究社会实际问题上，即社会研究者主要为定量统计方法的“消费者”而非“生产者”。作为“消费者”，学完这些课程后，我们一方面能够读懂、欣赏和评价别人在同行评议的刊物上发表的定量研究的文章；另一方面，也能在自己的研究中运用这些成熟的方法论技术。

上述两门课的内容，尽管在线性回归模型的内容上有少

量重复,但各有侧重。“社会科学里的统计学”从介绍最基本的社会研究方法论和统计学原理开始,到多元线性回归模型结束,内容涵盖了描述性统计的基本方法、统计推论的原理、假设检验、列联表分析、方差和协方差分析、简单线性回归模型、多元线性回归模型,以及线性回归模型的假设和模型诊断。“社会科学中的定量分析”则介绍在经典线性回归模型的假设不成立的情况下的一些模型和方法,将重点放在因变量为定类数据的分析模型上,包括两分类的 logistic 回归模型、多分类 logistic 回归模型、定序 logistic 回归模型、条件 logistic 回归模型、多维列联表的对数线性和对数乘积模型、有关删节数据的模型、纵贯数据的分析模型,包括追踪研究和事件史的分析方法。这些模型在社会科学研究中有着更加广泛的应用。

修读过这些课程的香港科技大学的研究生,一直鼓励和支持我将两门课的讲稿结集出版,并帮助我将原来的英文课程讲稿译成了中文。但是,由于种种原因,这两本书拖了多年还没有完成。世界著名的出版社 SAGE 的“定量社会科学研究”丛书闻名遐迩,每本书都写得通俗易懂,与我的教学理念是相通的。当格致出版社向我提出从这套丛书中精选一批翻译,以飨中文读者时,我非常支持这个想法,因为这从某种程度上弥补了我的教科书未能出版的遗憾。

翻译是一件吃力不讨好的事。不但要有对中英文两种语言的精准把握能力,还要有对实质内容有较深的理解能力,而这套丛书涵盖的又恰恰是社会科学中技术性非常强的内容,只有语言能力是远远不能胜任的。在短短的一年时间里,我们组织了来自中国内地及香港、台湾地区的二十几位

研究生参与了这项工程，他们当时大部分是香港科技大学的硕士和博士研究生，受过严格的社会科学统计方法的训练，也有来自美国等地对定量研究感兴趣的博士研究生。他们是香港科技大学社会科学部博士研究生蒋勤、李骏、盛智明、叶华、张卓妮、郑冰岛，硕士研究生贺光烨、李兰、林毓玲、肖东亮、辛济云、於嘉、余珊珊，应用社会经济研究中心研究员李俊秀；香港大学教育学院博士研究生洪岩璧；北京大学社会学系博士研究生李丁、赵亮员；中国人民大学人口学系讲师巫锡炜；中国台湾"中央"研究院社会学所助理研究员林宗弘；南京师范大学心理学系副教授陈陈；美国北卡罗来纳大学教堂山分校社会学系博士候选人姜念涛；美国加州大学洛杉矶分校社会学系博士研究生宋曦；哈佛大学社会学系博士研究生郭茂灿和周韵。

参与这项工作的许多译者目前都已经毕业，大多成为中国内地以及香港、台湾等地区高校和研究机构定量社会科学方法教学和研究的骨干。不少译者反映，翻译工作本身也是他们学习相关定量方法的有效途径。鉴于此，当格致出版社和 SAGE 出版社决定在"格致方法 · 定量研究系列"丛书中推出另外一批新品种时，香港科技大学社会科学部的研究生仍然是主要力量。特别值得一提的是，香港科技大学应用社会经济研究中心与上海大学社会学院自 2012 年夏季开始，在上海（夏季）和广州南沙（冬季）联合举办《应用社会科学研究方法研修班》，至今已经成功举办三届。研修课程设计体现"化整为零、循序渐进、中文教学、学以致用"的方针，吸引了一大批有志于从事定量社会科学研究的博士生和青年学者。他们中的不少人也参与了翻译和校对的工作。他们在

繁忙的学习和研究之余，历经近两年的时间，完成了三十多本新书的翻译任务，使得“格致方法·定量研究系列”丛书更加丰富和完善。他们是：东南大学社会学系副教授洪岩璧，香港科技大学社会科学部博士研究生贺光烨、李忠路、王佳、王彦蓉、许多多，硕士研究生范新光、缪佳、武玲蔚、臧晓露、曾东林，原硕士研究生李兰，密歇根大学社会学系博士研究生王骁，纽约大学社会学系博士研究生温芳琪，牛津大学社会学系研究生周穆之，上海大学社会学院博士研究生陈伟等。

陈伟、范新光、贺光烨、洪岩璧、李忠路、缪佳、王佳、武玲蔚、许多多、曾东林、周穆之，以及香港科技大学社会科学部硕士研究生陈佳莹，上海大学社会学院硕士研究生梁海祥还协助主编做了大量的审校工作。格致出版社编辑高璇不遗余力地推动本丛书的继续出版，并且在这个过程中表现出极大的耐心和高度的专业精神。对他们付出的劳动，我在此致以诚挚的谢意。当然，每本书因本身内容和译者的行文风格有所差异，校对未免挂一漏万，术语的标准译法方面还有很大的改进空间。我们欢迎广大读者提出建设性的批评和建议，以便再版时修订。

我们希望本丛书的持续出版，能为进一步提升国内社会科学定量教学和研究水平作出一点贡献。

吴晓刚

于香港九龙清水湾

序

经典回归模型的表达式很直接，以最常见的三变量多元回归为例，Y 是因变量，X_1 和 X_2 是自变量，可把方程 1 写成如下形式：

$$Y = a + b_1 X_1 + b_2 X_2 + e \qquad [1]$$

这里的 a、b_1 和 b_2 都是待估计的总体参数，e 是残差。假设满足必要的假定，那么普通最小二乘法(OLS)估计就是最佳的线性无偏估计。这个 OLS 方程是加叠性的，各项共同解释 Y，没有自变量相乘的项。假设自变量相乘产生乘积项 $X_1 X_2$，那么就可改写上面的方程为方程 2：

$$Y = a + b_1 X_1 + b_2 X_2 + b_3 (X_1 X_2) + e \qquad [2]$$

现在，我们有了一个既包括加叠性变量，又包括非加叠变量的方程。而乘积项或乘法的项目($X_1 X_2$)也被称为“交互作用项”，它的系数 b_3 估计交互作用。问题是，它的含义是什么呢？

通常，社会科学专业一年级的研究生会觉得这个问题挺难回答。他们这么理解加叠模型(方程 1)：当 X_2 保持不变时，X_1 变化一个单位，会导致 Y 变化 b_1 个单位。但如何理解 b_3，

即方程 2 中非加叠项的系数？当 X_1 对 Y 的影响取决于 X_2 的取值时，就存在交互作用。b_3 估计的正是这种交互作用。请考虑这个例子：$Y=$ 个人对竞选运动的捐献额（美元），$X_1=$ 收入（美元），$X_2=$ 教育程度（0 = 未上大学，1 = 大学及以上）。如果研究者认为对受过大学教育的人而言，其收入对其捐献额影响更大（相比未受过大学教育者收入的作用），那么研究者会偏向于使用包含交互作用的模型设定（方程 2），而非方程 1。

很明显，交互作用假设——一个变量的影响取决于另一个变量的取值——应当需要检验。回顾已发表的研究，我们发现事实并非如此（很多研究不检验变量间的交互作用），而我认为其原因在于交互作用这个想法较难捕捉，因此本书非常有意义。Jaccard 博士和 Turrisi 博士使理解和估计回归模型中的交互作用变得非常清楚。比如，他们强调，仅仅通过调节变量在不同组之间的作用差异来检验交互作用是不够的，如在上例中，仅仅分别观察上大学和没上大学两组人的收入在（分别的）回归方程中对政治捐献的影响是不够的。进而，他们提供了一个常常被忽略的见解：真正的度量来自数据本身，而不是测度。因此，要用回归检验交互作用，重要的是测度接近数据内在定距属性的程度。同时，本书也有许多新主题，比如聚类数据的交互作用以及随机系数模型。事实上，仔细阅读参考文献部分，我们能发现，1990 年之后出版的文献有 30 多种。

研究实践者必须利用工具去假定、理解和估计交互作用。仔细学习此书能使研究者在回归分析中达到上述要求。若要在其他或更高级的研究情境下研究交互作用，读者可参考 Jaccard 教授的其他著作。

迈克尔·S. 刘易斯-贝克

目 录

第1章

导　论

许多社会科学理论框架都关注因果模型。这些模型试图设定一个或多个自变量对一个或多个因变量或结果变量的影响。简单地说,因果模型存在六种可能关系,如图 1.1 所示。直接因果是一个变量 X 为另一个变量 Y 的直接原因。在理论系统语境下,X 就是 Y 的直接决定因素。间接因果是 X 对 Y 有因果作用,但 X 对 Y 产生的作用只能通过另一个变量 Z 来实现;虚假因果就是 X 和 Y 相关,但这种相关性完全是因为存在一个共同原因 Z,即 X 和 Y 不存在真正的因果关系;互为因果是指 X 对 Y 有因果影响,Y 对 X 也有因果影响;待定因

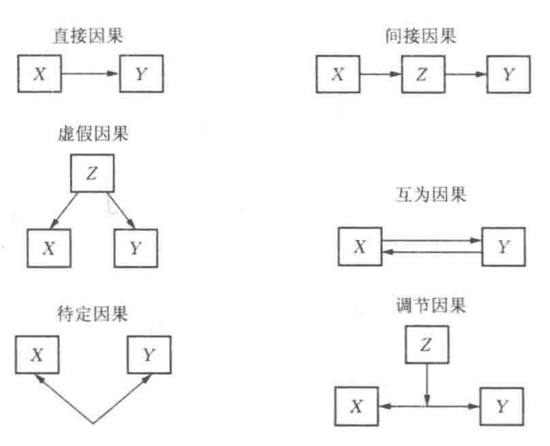


图 1.1　因果关系的类型

果是指 X 和 Y 相关，但这种关联的来源并未得到确认；最后，调节因果指 X 和 Y 之间的关系受到另外一个变量 Z 的调节作用。换句话说，X 和 Y 的关系是变动的，这种变动取决于 Z 的取值。

本书考察的是调节关联的统计分析，关注至少有一个自变量（或解释变量）是连续变量的情况。调节关系常被称为"交互作用"，尽管不同统计模型对交互作用的精确概念化有所不同。我们的关注点在于分析多元回归中的调节关联。在使用多元回归方法进行涉及连续变量的调节关联分析时，存在一些困惑需要厘清。

就如何最好地检验这些模型，现有文献充斥着大量相互矛盾的建议和忠告。本书最主要的目的就是梳理有关此主题的零散文献，说明进行包含连续变量调节关系的分析所涉及的中心议题。我们的目标是为那些具备多元回归操作知识的社会科学研究者展现一个合理的、易操作的导论。

在这一章中，我们会简单地回顾社会科学研究中交互作用分析的关键概念。首先考察一般意义上的交互作用概念，然后分析简单主效应和交互对比的概念。我们将回顾基本的多元回归程序，为下面各章节提供一个框架，包括虚拟变量的使用、简单变换对回归系数的影响以及预测值的计算。

第1节 | 交互的概念

如上面指出的，在社会科学中，许多方法都对交互作用进行了概念化，并且在哪种方法是最佳方法这一点上，存在很大的争议。一个流行的概念化交互作用的流派是把它当做调节关系。这一视角可用一个三变量系统进行展示，在此系统中，第一个变量被视为结果变量，第二个变量被视为自变量，第三个变量被视为调节变量，并且，结果变量被认为受到自变量的影响。而当自变量对因变量的影响取决于第三个变量的取值时（这一变量又称调节变量），就存在交互作用。例如，受教育年限对收入的影响可能因民族的不同而不同。教育对某些民族收入的影响要比另一些民族大。在调节变量框架中，收入是结果变量，受教育年限是自变量，而民族就是调节变量。再如，社会阶级对人们使用保健站频率的影响可能因性别的不同而不同。在此个案中，某人使用保健站的频率是结果变量或者因变量，社会阶级是自变量，而性别则是调节变量。性别被认为能够“调节”社会阶级对保健站使用情况的影响。

利用调节变量的方法进行交互分析要求理论家确定一个调节变量以及所谓的关键自变量。关键自变量对因变量的影响被认为根据调节自变量的某种函数关系而变化。大

部分正式研究都很容易确定一个解释变量作为“调节者”。例如,研究者想要确定情绪低落的临床治疗是否对男性比对女性更有效。在此个案中,很明显性别是调节变量,而实施治疗与否是关键自变量。

当一个理论家的调节变量是另一个理论家的关键自变量(或反之)时,那么就出现问题了。例如,一个研究产品质量与产品选择的消费心理学家或许对产品质量如何影响产品购买决定,而这又如何受到产品价格的调节作用感兴趣。相反,一个市场研究者使用消费心理学家的实验范式,但他可能对产品价格影响产品购买行为,而这种影响又如何受到产品质量的调节作用更感兴趣。在这两个例子中,研究者只是依据实践和理论方面的指导来确定调节变量,并没有设定一个比另一个好。而且,从统计上讲,这两种概念化方式下的交互分析会得到同样的统计结果。这两种设定方式代表的是针对同一现象的不同视角,不同视角引导研究者强调数据的不同方面。

有时,理论家也不确定哪个变量应当具有调节者的地位。假设研究者检验的是性别和民族对堕胎态度的影响,他/她想要确定的是对堕胎态度的性别差异如何作为民族的一个函数而变动。在此情况下,性别是关键自变量,民族是调节变量。研究者或许又会对堕胎态度的民族差异作为性别的一个函数感兴趣,在此情况下,民族是关键自变量,性别是调节变量。没什么能阻止研究者在这种情况下使用任一方案来刻画数据,这只不过涉及我们到底对什么感兴趣的问题。在第 4 章中,还会深入讨论此问题。

Jaccard 和 Dodge 讨论了其他概念化统计交互作用的方

法。调节方法在大量的研究领域被援引,甚至研究者有时并未意识到他们在这么做。在应用性研究情境下给交互作用灌输实质含义有些困难,大部分成功者最终都回到了调节变量的框架下。既然这一框架如此流行且易于理解,本书也将采用这一框架。第 4 章回顾了其他思考交互作用的框架。

第 2 节 | 简单效应与交互比较

在调节变量分析中有两个重要概念，简单效应(亦称“简单主效应”)和单自由度交互比较。展示这些概念的最好途径是使用定性解释变量的析因设计。假设某社会科学家找到 300 名已婚人士，一半为男性，一半为女性(但男性与女性均为非配偶关系)。

表 1.1 性别与宗教对堕胎态度的影响

	天主教	新 教	犹太教
女 性	5.0	6.0	7.0
男 性	3.0	3.0	3.0

在这些人中，1/3 是天主教徒，1/3 是新教徒，1/3 是犹太教徒。这形成了一个 2×3 的由性别和宗教交叉而成的析因设计。令研究者感兴趣的是性别和宗教如何与堕胎态度联系在一起，堕胎态度用 0 到 10 的尺度来测量，分数越高代表越赞同。每个子群体的均值列在表 1.1 中。假设在这个特殊的例子中，研究者决定把宗教信仰作为调节变量，而把性别作为关键自变量。

简单效应

研究者利用这些数据可以处理很多问题。一个常见问

题是,关键自变量(性别)是否在调节变量的每一个类别上都有不同的影响?要回答这个问题,就需要单独比较天主教男性和女性的均值以及新教徒的均值,最后比较犹太教徒男性和女性的差异。研究者在调节变量的三个类别中进行了三次显著性检验。这些比较通常被称为“简单效应”或“简单主效应”。它们关注的是关键自变量在给定的调节变量类别中对结果变量的作用。以表 1.1 中的数据为例,天主教徒的性别效应是对男性与女性态度差异 (5.0－3.0＝2.0) 的显著性检验。

另一种看待简单效应的方法是把它们当做条件效应。一个自变量对结果变量的简单效应的条件是调节变量取某特定值(例如,调节变量取值为“犹太教”)。

交互比较

简单主效应往往是研究者关心的问题,但它们与交互作用关系不大(尽管有些研究者认为它们有影响,稍后我们会讨论这一点)。对调节变量框架下的交互作用而言,关键自变量对结果变量的作用必须因调节变量取值的不同而变动。例如,天主教徒的性别差异必然与新教徒的性别差异不同,天主教徒的性别差异和犹太教徒的性别差异不同,新教徒的性别差异与犹太教徒的性别差异也不同。从表 1.1 的数据来看,天主教徒的性别差异是女性均值减去男性均值,即 5.0－3.0＝2.0,新教徒的性别差异是 6.0－3.0＝3.0,那么天主教徒的性别差异(2.0)就比新教徒的性别差异(3.0)小,这表明存在交互效应,即性别对堕胎态度的影响取决于宗教类别。

这一效应可以通过计算两个均值差之间的差 3.0－2.0＝1.0来反映。这一交互作用参数不为 0，表明确实存在交互作用。当然，非零取值也可能只是反映抽样误差，所以有必要对估计参数进行正式的显著性检验。

上文展示了所谓的单自由度交互比较。之所以称之为“交互作用”，是因为它明确地比较了关键自变量在调节变量一个类别中与在另一个类别中，对因变量的相应作用。天主教徒的性别差异(2.0)正式地与新教徒的性别差异(3.0)形成了对比。这种比较的统计检验在分子上只有一个自由度，因此就被称为“单自由度的交互比较”。关于这种比较的更多讨论，请参考 Jaccard(1996)。

单自由度的交互比较是交互分析的核心，它们代表了交互作用的关键检验。它们与整体交互检验不同，后者涉及一个以上的自由度。整体交互显著性检验是交互作用的总体检验，关注自变量和调节变量整体的而非其子群体的交互作用。在表 1.1 的析因设计中，整体检验关注的是性别和宗教之间总的交互作用，此例中自由度(df)为 2。相对而言，单自由度交互作用关注总样本中的 2×2 子表。在实践中，研究者很少满足于只在总体层面作出陈述。通常，更多重要的问题都会采用单自由度的交互比较。

现在，我们明确说明了为何简单主效应不能解释统计交互作用的机制。交互比较正式地比较了一个自变量在调节变量的一个类别上与另一个类别上对因变量影响的差异。相对而言，简单主效应并没有进行这种比较。简单主效应关注的只是调节变量在某一个类别上时，自变量是否在此特定类别中对因变量有影响。例如，对天主教徒而言，是否存在

性别差异？简单效应的显著性检验并没有比较其他群体的影响(例如，它没有比较新教徒和犹太教徒的效应)，它不能解决统计交互问题。

我们用另一种方法在一个相关的例子中来说明这一点。假设有两个变量，X 和 Y，男性 X 与 Y 的相关系数是 0.24，女性的是 0.22。假设对男性而言，这一相关系数是显著的($p < 0.05$)，但对女性是不显著的($p > 0.05$)。每组中简单效应的显著性检验是类似的。那我们能否从这些检验中得出结论说，X 和 Y 的相关性对男性来说比女性更强？当然不能。即使相关性在一组中统计显著，而在另一组不显著，我们也只能说，如果直接检验这两个相关性，那么它们存在差异。这种两个相关性差异的检验(在此例中并不统计显著)与调节变量框架下的交互作用的检验是类似的。

概言之，后面的章节对交互分析的讨论会考虑简单效应(也被称为“条件效应”)、单自由度交互比较和整体交互检验。这几种检验都具有潜在的理论意义，在使用交互模型的多元回归方法中都会出现。我们会使用定性变量和简单析因设计来展示这些概念，也会在包含连续性解释变量的回归模型中考察相应的情况。

第3节 | 多元回归分析的回顾

本书假定读者熟悉多元回归的基础知识，本节的目的是介绍一些术语，并建构一个讨论问题的参考框架。多元回归的导论可以参考 Berry 和 Feldman(1985)，或 Schroeder、Sjoquist 和 Stephan(1986)的著作。与传统方法一样，我们会使用希腊符号与字母来代表总体系数和总体数据，用阿拉伯数字来代表样本估计系数和样本数据。

线性模型

请考虑三个连续变量的例子，研究者关心两个自变量 X_1 和 X_2 对一个因变量 Y 的影响。分析样本数据主要采取的是最小二乘法方程：

$$Y = a + b_1 X_1 + b_2 X_2 + e$$

在这里，a 代表最小二乘法估计的总体截距，b_1 和 b_2 是最小二乘法估计中的 X_1 和 X_2 的总体回归系数，e 是残差。此方法要求总体数据结构满足如下条件，才能应用普通最小二乘法(OLS)并进行严格意义上的推论检验：(1)待检验线性模型适用于总体所有成员的真实模型；(2)总体残差独立分布，

均值为0;(3)解释变量本质上是固定的,且方差为正;(4)样本数据矩阵的阶等于列的数目,并少于观察值数量(例如,没有完全的多重共线性);(5)残差在一组给定的 X 取值下是正态分布的,且方差等于任何其他固定 X 取值下残差的方差(即不存在异方差性)。当以上假定都得到满足时,OLS估计就被称为"最佳线性无偏估计",因为它是线性的、无偏的,且具有一组无偏估计值中最小的方差。放松某些假定只会给推论结果和参数估计带来较小的影响,而违反某些假设则会带来问题。例如,尽管许多研究实践依赖于解释变量是随机的、非固定的,但OLS依然是有效的,只需研究者假定随机解释变量是以观察到的 X 为条件的一个真实样本。第4章会深入讨论违反假定的情况。

样本的复相关系数 R 是样本数据中模型整体拟合程度的指标。回归系数则通常被理解为控制方程中其他 X 变量时,某 X 变量对 Y 的影响。具体而言,某个特定的 b 代表当所有其他 X 都保持不变时,X 变化一个单位,Y 可能变化的单位。如果变量被标准化,那么截距一直为0,而 b 就代表了标准化回归系数;对系数的理解是一样的,只不过单位变成了标准分数。[1]

因此,如果解释变量 X 的标准化回归系数 b 为1.5,就表示 X 每增加一个标准分,Y 将会变动1.5个标准分。有些社会科学家在多元回归中更倾向于使用标准化回归系数,而不是非标准化回归系数,因为标准化之后,所有变量都被认为具有共同的度量,更容易实质性地比较不同自变量系数的大小。第4章会讨论这一问题的细节。

请考虑下面这个例子。某个社会学家对此问题感兴趣:

通过六类关系的满意或不满意来预测人们对自己婚姻关系的满意程度。共有 339 个人给出其对婚姻的满意或不满意分数，为 11 分制，从－5 到＋5，表示从非常不满意到非常满意（分数越高，满意度越高）。另外，他们还给出了对其生活中六个方面关系的满意程度（同样是 11 分制）：交往数量、情感表达方式、情感支持数量、共同爱好层次、共处时间以及争议解决方式。调查者进行多元回归分析，用六个因子对婚姻满意程度进行回归。表 1.2 展示了 SPSS 软件的电脑输出结果。

复相关系数平方是 0.663，这表示总体婚姻满意度的方差能被六因素线性组合所解释的比例。虚无假设指的是，总体的复相关系数为 0。通过 F 检验来检验这一假设，结果在表的上方。F 值是统计显著的 $[F(6,332)=108.70]$，这就拒绝了虚无假设。标准化和非标准化的回归系数在表 1.2 的下方。如上所述，非标准化回归系数反映了其他 X 变量不变时，X 变动一个单位导致总体婚姻满意度的变动。例如，情感支持满意度每增加一个单位，总体婚姻关系满意度会增加 0.307 个单位。相比之下，共处时间每增加一个单位，只能使总体满意度增加 0.005 个单位，当然，这都是在控制其他变量不变的情况下。标准化回归系数的理解是类似的，不过不是在原始分数，而是在标准分数的意义上。

每个非标准化回归系数都有一个标准误（在“SE/标准误”那一栏）。这些统计数字代表估计总体回归系数时抽样误差的大小。具体而言，估计的标准误表明针对随机样本大小为 N 的所有抽样，样本估计值与总体真实参数值之间的平均偏离水平。当其他条件不变时，标准误越大，抽样误差就

越大，我们对样本估计也就越没有信心。

标有“t”的那栏代表虚无假设的 t 检验，即某一特定的回归系数等于 0。t 值等于回归系数除以估计值的标准误。标有“Sig.”的那栏表示的是 t 检验的 p 值。在上例中，除（婚姻满意度与）共处时间满意度的相关性之外，其他的回归系数都是统计显著的（$p < 0.05$）。

最后，每一解释变量与标杆（指因变量）的相关性还可通过解释变量与因变量的零阶相关性（见“Zero-order”栏），以及解释变量和因变量的偏相关性来展示（控制住其他解释变量的影响，见“Part”栏）。前一统计值（零阶相关系数）的平方反映的是当所有其他的满意因素都允许自由变动时，某解释变量能够解释的总体婚姻满意度的比例。而后一个统计值（偏相关系数）的平方反映的是当其他变量都被控制时，这一变量单独能够解释的因变量方差的比例。例如，情感支持满意度能够解释总体婚姻满意度方差的比例是 $(100) \times (0.72) \times (0.72) = 51.5\%$，而它能够单独解释的因变量方差的比例是 $(100) \times (0.24) \times (0.24) = 5.8\%$（其他部分由另外五个解释变量所解释）。[2]

层级回归

研究者通常进行多元层级回归。在此情况下，研究者对加入一个或更多解释变量能否显著增加对因变量的解释力感兴趣。通常额外解释的方差是这样计算的：用原方程的复相关系数平方减去扩展方程的复相关系数平方。两个复相关系数平方的差异就是额外解释变量所能解释的因变量的

方差。

例如,如果差异等于 0.10,那么通过引入新的解释变量可额外解释因变量方差的 10%。可通过下面这个方程检验复相关系数平方在总体中为 0 的虚无假设:

$$F=\frac{(R_2^2-R_1^2)/(k_2-k_1)}{(1-R_2^2)/(N-k_2-1)} \qquad [1.1]$$

在这里,R_2 代表的是扩展方程的复相关系数,而 R_1 则是原方程的复相关系数;k_2 是扩展方程的解释变量数目,而 k_1 是原方程的解释变量数目;N 则是总样本量。F 值服从 k_2-k_1 和 $N-k_2-1$ 个自由度的 F 分布。

如上例,层级检验常被应用在向基本模型中增加一个解释变量的这种情况。然而,有时它会扮演从方程中删减变量的角色。假设回归方程有六个解释变量,我们希望判断去掉其中两个变量对复相关系数的影响。两个方程的复相关系数平方的变化反映了去掉两个变量时,方程能解释的因变量方差比例的变化。我们利用方程 1.1 来考察去掉两个变量后,模型所能解释方差的变化的统计显著性。这里,R_2 表示解释变量较多的模型的复相关系数,而 R_1 代表解释变量较少的模型的复相关系数。

分类解释变量和虚拟变量

回归分析通常包含分类变量作为解释变量,如性别、民族、宗教信仰。这些变量在方程中用虚拟变量来表示。分析者为确定属于哪一群组而生成一个虚拟变量。以性别变量为例,生成一个虚拟变量,然后所有男性赋值为 1,所有女性

表 1.2 SPSS 多元回归输出结果

模 型	R	R^2	调节的 R^2	测得 SE
1	0.814	0.663	0.457	0.74094

	平方和	df	平方均值	F	Sig.
回归系数	357.897	6	59.650	108.698	0.000
残 差	182.191	332	0.549		
总 计	540.088	338			

模 型	非标准化回归系数		标准化回归系数	t	Sig.	B 的 95%置信区间		相关系数	
	B	SE	β			低端	高端	Zero-order	Part
(常数项)	1.010	0.127		7.980	0.000	0.761	1.259		
争议解决	0.121	0.025	0.191	4.880	0.000	0.072	0.170	0.578	0.155
共处时间	0.005	0.023	0.008	0.220	0.823	−0.044	0.054	0.353	0.007
情感表达	0.151	0.039	0.173	3.910	0.000	0.075	0.227	0.629	0.210
共同爱好	0.153	0.034	0.183	4.460	0.001	0.086	0.220	0.550	0.142
交往数量	0.125	0.037	0.147	3.330	0.000	0.058	0.192	0.625	0.106
情感支持	0.307	0.041	0.350	7.560	0.000	0.227	0.387	0.724	0.241

赋值为0。这种赋值方法称为“虚拟编码”，即所有属于某一群组的个案赋值为1，所有不属于这一群组的个案赋值为0。当一个定性变量有两个以上的类别时，必须设定一个以上的虚拟变量来表示不同的群组。一般而言，我们需要 $m-1$ 个虚拟变量，这里的 m 指的是一个变量的类别。

假设有一个解释变量——党派，有三个取值——民主党、共和党或独立党。在此例中，需要 $3-1=2$ 个虚拟变量来代表党派。第一个虚拟变量是 D_{D}，所有的民主党党员都赋值为1，其他人都赋值为0。第二个虚拟变量是 D_{R}，共和党党员赋值为1，其他人都赋值为0。尽管可以产生第三个虚拟变量来代表独立党，并且独立党党员赋值为1，其他人赋值为0，但这一变量与另外两个虚拟变量放在一起是冗余的。一旦知道某人是否为民主党党员，且知道某人是否为共和党党员（通过第一个和第二个虚拟变量的值），根据定义就自然知道此人是否为独立党党员。这一点背后的逻辑对性别变量来说更明显。生成一个虚拟变量来区分这两个群体，男性赋值为1，女性赋值为0。如果产生第二个虚拟变量，女性赋值为0，男性赋值为1，那么这个虚拟变量就与第一个虚拟变量完全负相关，因此是冗余的。通过虚拟编码，那些在任一虚拟变量上都未被赋值为1的个案组成了参照组。在上例中，性别的参照组是女性；而在党派例子中，参照组是独立党。从统计角度来讲，可任意选择一组作为参照组。

假设用对堕胎态度的一个测度对宗教类别这一分类变量进行回归。态度变量的取值是0到100，分数越高表明越支持。宗教变量有三组，即天主教、新教和犹太教，它由两个虚拟变量来表示，一个虚拟变量是所有天主教徒赋值为1，其

他教派为 0(D_C);另一个是所有的新教徒赋值为 1,而其他人赋值为0(D_P)。犹太教徒是参照组。假设回归分析的结果是复相关系数平方 0.30,这表明宗教能够解释堕胎态度 30%的方差。针对复相关系数平方的显著性检验——F 检验,检验的虚无假设是宗教对堕胎态度没有影响,或三个宗教群体的堕胎态度均值无差异。两个虚拟变量的非标准化回归系数很容易理解,每个系数都代表了一个平均的差异。具体而言,系数代表这个取值为 1 的组的平均得分与参照组平均得分之差。假设 D_C 的回归系数是−1.0,这表明天主教徒态度分数均值减去犹太教徒的均值是−1.0,或者说,平均而言,天主教徒的堕胎态度分数比犹太教徒的堕胎态度分数低 1.0 个单位。D_P 的回归系数是−0.5,这表明,平均而言,新教徒的堕胎态度分数比犹太教徒的堕胎态度分数低 0.5 个单位。这些回归系数的显著性检验,检验的是这些平均差异的显著程度。

如果重复上面的分析,但加入一个连续变量到回归方程中,那么虚拟变量的回归系数依然表示平均差异,不过这是模型中其他变量保持不变的情况下的平均差异(例如,上述连续变量)。例如,在方程中加入社会阶级这一测度,则 D_C 的回归系数变成了−0.80,这就是在控制社会阶级变量时,天主教徒和犹太教徒堕胎态度的平均差异。

还有许多给虚拟变量赋值的方法。如上所述,依靠 1 和 0 的赋值方法被称为"虚拟编码"。Hardy 讨论了不同编码方案的逻辑(Hardy, 1993)。虚拟编码很有用,因为它很好地与交互作用的调节变量概念匹配,在下面的章节中会分析这一点。

多元回归中的预测值

假设一组数据分析的是性别和用保守主义—自由主义来测量的意识形态对一个环境组织提倡的清理河道运动的捐款数目(单位为美元)的影响。意识形态测度的取值范围是−3到+3,0代表中立点,负值代表程度更高的保守主义,而正值代表程度更高的自由主义;性别用虚拟变量来代表,男性取值为1,女性取值为0;结果变量的单位是美元。假设分析得到下面这个回归方程:

$$Y' = 10.00 + (-5.50)Gender + 1.00Ideology \quad [1.2]$$

这里的Y'是预测的个体捐款金额。给定一个解释变量的取值,把这个值代入方程中,就可计算Y的预测值。例如,对一个意识形态分数为+2的男性来说,其捐款金额预测值就是:

$$Y' = 10.00 + (-5.50) \times 1 + 1.00 \times (2) = 6.50$$

对那些意识形态分数为−2的男性,其捐款金额的预测值又是多少呢?通过代换,我们得到:

$$Y' = 10.00 + (-5.50) \times 1 + 1.00 \times (-2) = 2.50$$

在后面的分析中,我们将会用到这些预测值。

解释变量的变换

通过在回归之前对解释变量进行代数处理,可使回归系

数直接反映那些有理论意义的参数。后面的章节会揭示这么做的好处，但在此，我们会阐释基本的逻辑。上例中，假定在回归分析之前，把意识形态变量减去一个常数 1。原先尺度的取值范围是−3 到+3，新尺度的取值范围变成了−4 到+2，每个分数都往下平移一个单位。运用这种变换分数的回归分析结果如下：

$$Y' = 11.00 + (-5.50)Gender + 1.00Ideology_{t}$$

请注意，只有一个参数受到这种转换的影响，即截距，而其他参数与原先的分析结果并无二致。截距表示性别为 0 且变换后意识形态分数也为 0 时的预测值。但在这里，变换后的意识形态分数 0，代表原先的意识形态分数为+1。第二个分析的截距应当等于原先的分析中意识形态分数为 1 的女性的预测值。事实正是如此。在原方程中：

$$Y' = 10.00 + (-5.50) \times 0 + 1.00 \times (1) = 11.00$$

这与第二次分析中的截距是相同的。为何要进行这种变换呢？几乎所有的电脑程序都不仅报告回归方程中的参数估计，还给出估计值的标准误和给定估计的置信区间。进行如上变换代表了一种简单的、计算任意解释变量组合的 Y 预测值置信区间的方法。简单变换每一个解释变量，通过增加或减去一个常数项，使变换后的 0 值代表原先尺度中研究者感兴趣的解释变量的取值。利用变换后的解释变量进行回归，方程的截距项就提供了自变量特定组合下预测的 Y 均值，且截距项的置信区间就是与此自变量组合相连的 Y 预测值的置信区间。

如果进行并非这种性质的变换，那么很多时候截距项就

不易理解，因为它反映了 Y 预测值的均值，但解释变量取值为 0 的情况并不存在，或者说超出了研究变量的取值范围。后面的章节还会用到这一变换策略及其变形，这是用来分离一系列显著性检验和置信区间的关键策略。

第 4 节 | 本书概览

本书余下的部分会讨论一系列与交互分析相关的议题。我们的关注点在于连续变量的分析以及连续变量和定性变量混在一起时的分析，因为精妙的交互分析同样出现在所有解释变量都是定性变量的情况下(例如，Cohen & Cohen, 1983)。第 2 章关注双向交互作用，并以传统的乘积项分析作为双向线性交互作用的分析工具。第 3 章考察三向交互作用。第 4 章关注各种已经整合到交互分析文献中的话题，并协助应用型研究者在运用交互模型时考虑一些重要问题。

第2章

双向交互作用

本章首先考察包含交互作用在内的多元回归模型设定的一般性问题。其次，考察两个解释变量都是连续变量时的双向交互作用。最后，我们讨论一个解释变量是定性的，而另一个是连续变量的情况。

第1节｜包含乘积项的回归模型

举一个有三个连续变量的例子:研究者对两个自变量(X 和 Z)对一个因变量(Y)的影响感兴趣。具体来说,假设研究者试图理解为何发生性行为时,有些青少年不采取避孕措施,而有些青少年却倾向于采取避孕措施。研究者获取了一个包含125名有性生活的女青年样本,并获取每个女青年采取避孕措施意愿的测度。这一测度包含一个评分尺度,一端是"绝对不愿采取避孕措施",另一端是"绝对希望采取避孕措施"。分数的取值是0到30,分数越高代表采取避孕措施的意愿越强。研究者假设有两组因素会影响这种意愿:第一组是个体对采取避孕措施的感受或者态度,即个体喜欢还是不喜欢避孕措施;第二组是感受到采取避孕措施的同侪压力。所有这些因素,态度和感受到的同侪压力,都用1到5的五分制尺度来衡量。对态度测度而言,分数越高,个体越倾向于采取避孕措施;对同侪压力测度而言,分数越低,代表越少的采取避孕措施的同侪压力,而分数越高则代表这种压力水平在上升。

如第1章指出的,检验加叠性(或"主效应")模型中 X 和 Z 解释 Y 的显著性,通常采取最小二乘法方程,总体模型如下:

$$Y = \alpha + \beta_1 X + \beta_2 Z + \varepsilon \qquad [2.1]$$

应用方程 2.1 涉及避孕意愿对态度和同侪压力进行回归。假设研究者对是否存在交互效应感兴趣。具体而言，研究者假设态度和意愿之间的关系受到同侪压力的调节作用：当同侪压力最小时，个体态度就会对意愿产生一个更强的影响；当同侪压力很大时，态度的影响力就小了。在此例中，关键自变量是使用避孕措施的态度（X），调节变量是同侪压力（Z）。

最常见的处理回归分析中交互作用的方法是使用乘积项。我们可以用简单的代数来展现乘积项的基本逻辑。如果态度对意愿的影响通过方程 2.1 中的 β_1 来反映，且 Z 被认为对态度对意愿的影响有调节作用，那么当 Z 变化时，β_1 的取值也会相应地变化。可用下面这个线性方程说明 β_1 为何可能是 Z 的函数：

$$\beta_1 = \alpha' + \beta_3 Z \qquad [2.2]$$

根据这个方程，Z 每变化一个单位，β_1 会被预计有 β_3 个单位的变化。现在，把右边部分代入方程 2.1 的 β_1 项，得到的结果是：

$$Y = \alpha + (\alpha' + \beta_3 Z)X + \beta_2 Z + \varepsilon$$

将 X 乘入，得到的结果是：

$$Y = \alpha + \alpha' X + \beta_3 ZX + \beta_2 Z + \varepsilon$$

那么，给系数一些新标签并调整顺序，我们就可以得到包含乘积项的交互模型：

$$Y = \alpha + \beta_1 X + \beta_2 Z + \beta_3 ZX + \varepsilon \qquad [2.3]$$

这一阐述过于简单，因为其他类型的交互模型会产生同样的方程，而其他类型的交互模型可以从不同的交互模型中演化而来。在此，只是给读者提供乘积项背后的逻辑来反映关键自变量对结果变量的作用被认为是调节变量线性函数的情况。

第 2 节 | 两个连续性解释变量

传统的回归策略

我们继续讨论那个用态度和同侪压力来预测(避孕)意愿的交互作用分析的例子。数据显示在表 2.1 中。这些数据是虚构的,用于反映非真实的系统性的属性,以揭示下面讨论中涉及的问题。出于教学的考虑,数据根据析因表,即每一个 X 和 Z 的组合给出 Y 相应的均值分数,这体现在表 2.2 中。这个表格本质上是 5×5 的析因设计。

多元回归中最流行的分析交互作用的策略是由 Cohen 等人(1983)提出的。这涉及一个乘积项,XZ,它被认为包含了交互作用,然后计算两个 R^2 的值,一个模型"只含主效应"(方程 2.1),另一个模型在主效应模型中加入乘积项,从而这组样本数据产生了一个有三个解释变量的方程:

$$Y = a + b_1 X + b_2 Z + b_3 XZ + e \qquad [2.4]$$

如果存在交互作用,那么两个 R^2 的差异应当是统计显著的(除非有类型 1 错误)。[3]可用方程 1.1 中的层级 F 检验对这种差异进行正式的显著性检验。

使用不包括乘积项的标准回归程序分析表 2.1 中的数据，得到两解释变量模型的复相关系数 R 是 0.90139。回归方程是：

$$Y = 8.0 + 3.0X + (-2.0)Z + e \qquad [2.5]$$

三项交互模型的复相关系数 R 是 0.96825，回归方程是：

$$Y = -1.0 + 6.0X + 1.0Z + (-1.0)XZ + e \qquad [2.6]$$

应用方程 1.1 得到下面这个结果：

$$F = \frac{(0.96825^2 - 0.90139^2)/(3-2)}{(1-0.96825^2)/(125-3-1)} = 242.26$$

表 2.1　交互作用的虚构数据

ID	Y	X	Z	ID	Y	X	Z	ID	Y	X	Z
1	3	1	1	21	3	1	5	41	5	2	4
2	4	1	1	22	4	1	5	42	6	2	4
3	5	1	1	23	5	1	5	43	7	2	4
4	6	1	1	24	6	1	5	44	8	2	4
5	7	1	1	25	7	1	5	45	9	2	4
6	3	1	2	26	8	2	1	46	4	2	5
7	4	1	2	27	9	2	1	47	5	2	5
8	5	1	2	28	10	2	1	48	6	2	5
9	6	1	2	29	11	2	1	49	7	2	5
10	7	1	2	30	12	2	1	50	8	2	5
11	3	1	3	31	7	2	2	51	13	3	1
12	4	1	3	32	8	2	2	52	14	3	1
13	5	1	3	33	9	2	2	53	15	3	1
14	6	1	3	34	10	2	2	54	16	3	1
15	7	1	3	35	11	2	2	55	17	3	1
16	3	1	4	36	6	2	3	56	11	3	2
17	4	1	4	37	7	2	3	57	12	3	2
18	5	1	4	38	8	2	3	58	13	3	2
19	6	1	4	39	9	2	3	59	14	3	2
20	7	1	4	40	10	2	3	60	15	3	2

续表

ID	Y	X	Z	ID	Y	X	Z	ID	Y	X	Z
61	9	3	3	83	17	4	2	105	27	5	1
62	10	3	3	84	18	4	2	106	19	5	2
63	11	3	3	85	19	4	2	107	20	5	2
64	12	3	3	86	12	4	3	108	21	5	2
65	13	3	3	87	13	4	3	109	22	5	2
66	7	3	4	88	14	4	3	110	23	5	2
67	8	3	4	89	15	4	3	111	15	5	3
68	9	3	4	90	16	4	3	112	16	5	3
69	10	3	4	91	9	4	4	113	17	5	3
70	11	3	4	92	10	4	4	114	18	5	3
71	5	3	5	93	11	4	4	115	19	5	3
72	6	3	5	94	12	4	4	116	11	5	4
73	7	3	5	95	13	4	4	117	12	5	4
74	8	3	5	96	6	4	5	118	13	5	4
75	9	3	5	97	7	4	5	119	14	5	4
76	18	4	1	98	8	4	5	120	15	5	4
77	19	4	1	99	9	4	5	121	7	5	5
78	20	4	1	100	10	4	5	122	8	5	5
79	21	4	1	101	23	5	1	123	9	5	5
80	22	4	1	102	24	5	1	124	10	5	5
81	15	4	2	103	25	5	1	125	11	5	5
82	16	4	2	104	26	5	1				

表 2.2 作为 X 和 Z 函数的单元格均值

态度(X)	同侪压力(Z)				
	1	2	3	4	5
5	25	21	17	13	9
4	20	17	14	11	8
3	15	13	11	9	7
2	10	9	8	7	6
1	5	5	5	5	5

对 1 和 121 的自由度(F 检验)而言，F 值是统计显著的，

表明存在统计交互作用。层级 F 检验产生的 p 值与方程 2.4 中 b_3 的显著性检验，即 t 检验的 p 值应当相同。在此样本中，b_3 的 t 值是 15.56，而这个值的平方正是大家所看到的 F 值，即 $15.56 \times 15.56 = 242.26$。因此，与其执行层级检验来决定交互作用的统计显著性，不如简单地检验 b_3 的统计显著性。

下面挑选一些有关这一分析的重要问题进行讨论。我们用避孕措施的例子来讨论这些问题，随后再用新的例子来展示研究实践中的应用。

交互的形式

如上所述，简单的乘积项允许研究者检验是否存在调节关系。原则上来说，有一系列调节关系可用于刻画两个连续变量之间的交互作用。例如，X 对 Y 作用的斜率是 Z 的线性函数。这被称为“双线性交互作用”，表 2.2 中的关系正是如此。当同侪压力低至 1 时，意愿对态度的斜率是较高的，即 5.0。态度每变化一个单位，意愿会改变 5.0 个单位。当同侪压力增加时，这一斜率就降低了。例如，当同侪压力为 3 时，意愿对态度的斜率变化为 3.0，即态度每变化一个单位，意愿对态度的斜率只变化三个单位。每当同侪压力增加一个单位，这个斜率就降低一个单位。斜率变化和调节变量变化之间这种顺序的、单调的线性关系，正是传统乘积项方法下统计交互作用的精华。

其他类型的函数关系也是可能的。例如，当某人从低同侪压力向中同侪压力变动时，意愿—态度斜率的变化更大。

然而，当某人从中同侪压力向高同侪压力转变的时候，斜率变化的程度会稍稍平缓，直到它们达到一个最小的变化点。另外，意愿和态度之间的关系可能是非线性的。自变量和因变量关系的曲线的形状（而非斜率）可能随着同侪压力的变动而变动。涉及连续性解释变量的可能函数关系的数量是无限的。在此需要强调，我们只使用了一种函数形式，即双线性交互作用，如方程 2.4 中的简单乘积项所示。若使用传统的乘积项却发现交互作用统计不显著，那么就说明可能存在另一种函数形式（的交互作用），而并非一定不存在调节关系。第 4 章会讨论识别交互作用其他形式的策略以及检验它们的策略。在理想情况下，理论应当预测一个特定的函数形式，然后，数据分析应当建构适当的模型来检验函数的形式。本节余下的部分将分析局限于双线性交互作用。

理解乘积项的回归系数

方程 2.4 中所有的回归系数都需要解释。我们首先考虑 b_3，乘积项的回归系数。b_3 是单自由度的交互比较，传达了交互作用的本质信息。它表明 Z 增加一个单位，X 对 Y 作用的斜率所变化的单位数量。在我们的样本中，b_3 为 -1.0，这表示同侪压力每增加一个单位，态度对意愿作用的斜率就降低一个单位。这可以从表 2.2 中看出来。

表 2.2 的数据是虚构的，在给定 Z 值的情况下，X 对 Y 作用的斜率可通过对数据的观察而发现。然而，研究实践中很少有这种情况。当使用三个回归系数可计算任意给定的 Z 值时，X 对 Y 的作用只需一些代数计算。

当要分析 Y 和 X 的关系时，必须先设定 Z（同侪压力）的取值。让我们从 Z 的最小可能取值 1 开始。利用三项式回归方程，把 1 代入 Z 的位置。得到下列方程：

$$Y = -1.0 + 6.0X + (1.0)(1) + (-1.0)(X)(1) + e$$

重新安排方程右边的项目，合并同类项，X 仍然在右边：

$$Y = -1.0 + (1.0)(1) + 6.0X + (-1.0)(X)(1) + e$$

接着把 X 的相关项合并：

$$Y^{[4]} = -1.0 + (1.0)(1) + [6.0 + (-1.0)(1)](X) + e$$

得到：

$$Y = 0.0 + 5.0X + e$$

这一结果是指当 Z 取 1 时，Y 和 X 之间关系的线性方程。接下来进行同样的计算，把 Z 的最大值 5 代入，得到的方程是：

$$Y = -1.0 + 6.0X + (1.0)(5) + (-1.0)(X)(5) + e$$

进行上面同样的计算，得到下面的结果：

$$Y = 4.0 + 1.0X + e$$

这些计算揭示了 Y 和 X 之间的关系如何在 Z 的取值范围内变动：在 Z 值的低端（比如，1），X 一个单位的变动能使 Y 的预测值变动五个单位；而在 Z 值的高端（比如，5），X 一个单位的变动使 Y 的预测值变动一个单位。可用一个方程来计算在 Z 的每个取值上，X 对 Y 预测值的作用，它就是：

$$b_1 \text{ 在 } Z = b_1 + b_3 Z \qquad [2.7]$$

以我们的例子来说，当 Z 的取值为 4 时，X 对 Y 的作用斜率

如下：

$$b_1 \text{ 在 } 4 = 6.0 + (-1.0)(4) = 2.0$$

若计算在 Z 的每一个取值上，X 取值对 Y 作用的斜率，那么得到：

$$b_1 \text{ 在 } Z \text{ 取值为 } 1 = 5.0$$
$$b_1 \text{ 在 } Z \text{ 取值为 } 2 = 4.0$$
$$b_1 \text{ 在 } Z \text{ 取值为 } 3 = 3.0$$
$$b_1 \text{ 在 } Z \text{ 取值为 } 4 = 2.0$$
$$b_1 \text{ 在 } Z \text{ 取值为 } 5 = 1.0$$

请注意，Z 每增加一个单位，斜率取值就降低 1.0 个单位（−1.0），这正是 b_3 的值。另外，乘积项系数取值揭示了调节变量变动一个单位时，斜率随之变动的单位数量。

理解组成项的回归系数

有些社会科学家强调，在一个方程中加入乘积项会使得原组成项（X 和 Z）的回归系数较难理解。研究者注意到，相比那些只包含主效应方程的相应系数，加入乘积项往往会使这些系数发生剧烈变动。事实上，系数甚至会改变方向。这种“矛盾”使在包含乘积项模型中理解回归系数变得困难。

事实上，三项式方程中的回归系数需要我们去解释。上面提到的这些差异是因为两个方程估计的是不同的概念。在两项只含主效应的模型中，回归系数估计一个解释变量对因变量的影响，并控制其他解释变量的每一个层次。在两项式模型中，b_1 反映的是在每一个 Z 值上，X 变动一个单位，Y

随之变动的单位；而 b_2 反映的是在每个 X 值上，Z 变动一个单位，Y 变动的单位数。对包含乘积项的模型而言，X 和 Z 的回归系数反映了基于特定 Z 和 X 值条件下的关系：b_1 反映的是当 Z 等于 0 时，X 对 Y 的影响；b_2 反映的是当 X 等于 0 时，Z 对 Y 的影响。这两个模型中 b_1 和 b_2 的系数有差异是因为在“主效应”模型中，这两个系数反映在其他解释变量取平均水平时，X 和 Z 的“通常”关系；而在乘积项模型中，它们反映包含乘积项的其他解释变量特定取值时的条件关系。[5]

换句话说，在乘积项模型中，X 和 Z 的回归系数并非如传统认为的那样代表主效应。相反，这些系数代表的是简单效应。X 的系数估计的是 Z 为特定值（即 $Z = 0$）时，X 对 Y 的作用，而 Z 的系数估计的是 X 为特定值（即 $X = 0$）时，Z 对 Y 的影响。有些研究者错误地理解了这些系数，把它们当成主效应，实际上它们不是。我们必须牢记这些系数的本质。

仍以避孕研究为例，X 的系数代表的是当同侪压力 Z 为 0 时，态度 X 对避孕意愿 Y 的影响。系数为 6.0，因此当同侪压力对应的 Z 值为 0 时，态度变化一个单位，预计会使意愿变化六个单位。

然而，这种理解有一个问题。系数表明了当 $Z=0$ 时，X 对 Y 的影响大小，但同侪压力为 0 的取值是不存在的。Z 值的范围是 1 到 5，所以当 $Z = 0$ 时，态度对意愿的影响其实没有实质意义。解决该问题的一个方法是变化 Z，使 0 值有意义。例如，Z 的均值是 3.0，假设我们把每个样本的 Z 的分数都减去 3.0，这样，与原取值范围 1 到 5 不同，变化后的 Z 值（Z_t）的取值范围将是 −2 到 +2。相比 Z 的均值为 3.0，Z_t 的

均值将是 0。一个变量减去自身均值以形成新尺度的方法被称为“对中”。

假设我们对中 Z_t，重新计算乘积项 XZ_t，然后估计回归方程：

$$Y = a + b_1X + b_2Z_t + b_3XZ_t + e \qquad [2.8]$$

原来使用 Z 作为解释变量的方程是：

$$Y = 1.0 + 6.0X + (1.0)Z + (-1.0)XZ + e$$

现在，Z_t 取代了 Z 的位置，方程变为：

$$Y = 2.0 + 3.0X + 1.0Z_t + (-1.0)XZ_t + e$$

这两个分析有几个特点。首先，利用 X、Z 和 XZ 来预测 Y 的复相关系数平方及其显著性与利用 X、Z_t 和 XZ_t 是一致的。复相关系数(及其显著性检验)不随变换而变动。其次，b_3 的值和它的显著性检验同样不随变换而改变。两个交互作用分析体现了完全相同的特征。尽管如此，请注意，X 系数从原分析中的 6.0 变成了 3.0。在新的分析中，X 系数是当 $Z_t = 0$ 时，X 对 Y 影响作用的大小。在此分析中，Z_t 为 0 对应原来的 Z 为 3。所以这一系数实际上反映的是 $Z = 3$ 时，X 对 Y 的影响大小。换句话说，当同侪压力中等时，若 Z 值为 3，那么 X 对 Y 的影响大小反映为 Y 对 X 回归的斜率 3.0。通过变换 Z 值，即通过使 0 值有意义，与 X 相连的系数也变得有意义了。在新的分析中，b_1 就是当 Z 等于其样本均值时，X 对 Y 的预测作用大小。

现在给定交互模型中的一个 Z 值，我们用另外两种方法来计算 X 对 Y 的作用。第一个策略是利用方程 2.7，我们可

在此重复一下：

$$b_1 \text{ 在 } Z = b_1 + b_3 Z$$

例如，当 $Z = 2$ 时，b_1 的值是 $6.0 + (-1.0)(2) = 4.0$。[6]第二个策略是将 Z 进行变换，使 Z_t 为 0 时，原度量 Z 值正好是有用的那个值，从而得到 Z_t，然后再计算相应的回归方程。例如，如果把 2 代入每一个 Z 值中，Z_t 取值范围就变成了 -1 到 3，而 0 对应的就是原来度量中的 2。而在 Y 对 X 和 Z_t 的回归中，得到 b_1 的值为 4.0。后一种策略看上去很繁琐，但作为考察调节变量取某特定值时 X 对 Y 作用大小的方法，后面的章节会说明其优势。

显著性检验与置信区间

所有的计算机统计软件的多元回归程序通常都会提供模型估计的回归系数的标准误、显著性检验和置信区间。b_3 的显著性检验的虚无假设是指不存在双线性交互作用。如前所述，t 检验与层级 F 检验产生相同的 p 值，因此，当只在主效应模型中加入一个乘积项时，就不必再进行传统的层级检验。

对交互模型中的“主效应”项，回归系数的条件性本质同样适用于估计系数的标准误。当其他乘积项中的解释变量为 0 时，X 和 Z 在交互模型中所反映的估计标准是条件性的，反映了抽样误差。因此，方程 2.4 中 b_1 的标准误就是当 Z 为 0 时，Y 对 X 回归的系数的抽样误差。简单地讲，方程 2.4 中 b_2 的标准误估计值是当 X 为 0 时，Y 对 Z 回归的系数

的抽样误差。

在前面的章节中,我们展示了研究者如何通过方程 2.7 在给定 Z 值时,计算 Y 对 X 回归的斜率。同时,该方程也可计算这一系数的标准误。利用包含乘积项的方程,估计的标准误就是:

$$SE(b_1 \text{ 在 } Z) = [\mathrm{var}(b_1) + Z^2 \mathrm{var}(b_3) + 2Z\mathrm{cov}(b_1, b_3)]^{1/2} \qquad [2.9]$$

当 $\mathrm{var}(b_1)$是回归系数 b_1 的方差,而 $\mathrm{var}(b_3)$是回归系数 b_3 的方差时,$\mathrm{cov}(b_1, b_3)$就是回归系数 b_1 和 b_3 的协方差。方程右边的方差和协方差通过标准电脑程序输出得到,尽管默认选项并不输出这些值。给定 Z 值时,b_1 的显著性检验形式是 t 检验,即用系数除以它的标准误,亦即:

$$t = (b_1 \text{ 在 } Z)/SE(b_1 \text{ 在 } Z) \qquad [2.10]$$

此时,Z 是那个检验 X 对 Y 影响时的 Z 值。方程 2.10 的 t 值服从 t 分布,自由度是 $N-k-1$, k 就是在交互作用模型中的解释变量的数目(在此例中, $k=3$)。

一个更简单的策略是使用前面所讨论的变换策略,在任何给定 Z 值的条件下,计算相关的标准误、置信区间以及 b_1 的显著性检验。通过减去一个常数进行简单的 Z 值变换,使 Z_t 值为 0 对应原先我们感兴趣的 Z 值,然后用电脑程序让 Y 对 X、Z_t 和 XZ_t 进行回归。b_1 的取值就是当 Z_t 为 0 时,X 对 Y 作用的系数;b_1 的标准误、显著性检验和置信区间都是此条件性系数所需的参数。下文会展示此方法的应用。

总结一下,传统的交互模型涉及两个连续性解释变量时,是利用下面这个方程对样本数据进行检验的:

$$Y = a + b_1X + b_2Z + b_3XZ + e$$

b_3 在此模型中的系数是单一自由度的交互比较,表示调节变量 Z 变动一个单位,Y 对 X 回归的斜率会变动的单位数量。b_3 的显著性检验的虚无假设是不存在双线性的交互作用。系数 b_1 是当 $Z=0$ 时,X 对 Y 作用的大小。研究者或许会对给定 Z 值时,X 对 Y 的作用大小感兴趣,此时变换变量策略就有效。这一策略能产生与这些简单效应相关的标准误、显著性检验和置信区间。

多重共线性

有些研究者常会害怕包含乘积项的交互分析,因为乘积项与定义这一乘积项的组成变量高度相关。如果 XZ 与 X、Z 或两者同时高度相关,那么研究者就会担心多重共线性会使交互作用被低估。这种担忧通常是有误导性的。

如前所述,在对交互作用的显著性进行检验时,Z 减去一个常数,或 X 减去一个常数,或 X 和 Z 都减去一个常数,b_3 的显著性不变。尽管交互作用系数的值和 t 检验都不受变换的影响,但变换确实影响 XZ 和 Z 之间的相关性以及 XZ 和 X 之间的相关性。

变换有时会增加相关性,有时会降低相关性。如果 X 和 Z 都正态分布,那么对中 X 和 Z 后再形成乘积项,会使乘积项与 X 和 Z 都不相关(参见 Cronbach, 1987)。尽管如此,b_3 的显著性检验和 b_3 的置信区间与变换前是相同的,而 XZ 与其组成部分的相关性却相当不同[7]。

在此,并不需要考虑这一结果的统计学原理,感兴趣的

研究者可参考 Friedrich(1982)和 Cronbach(1987)的著作。高度共线性的主要观点是,乘积项及其组成项之间的共线性并不会使交互分析有问题,除非共线性高到足以破坏电脑代数运算的设计,使统计软件无法分离相关的标准误。如果是这种情况(可从电脑给出的错误提示中得到证实),研究者可简单地对中 X 和 Z(减去各自的均值),然后再分析。在大部分情况下,这会大大降低共线性,从而解决电脑程序上的问题。

尽管 XZ 和 X 以及 XZ 和 Z 的共线性没有问题,但这并不代表 X 和 Z 的共线性也没有问题。X 和 Z 的高度共线性会导致严重的后果。

交互作用的优点

交互作用的优点可通过各种统计方法进行评估,或通过非标准化作用大小的测度,或通过标准化作用大小的测度。最流行的非标准化指标是 b_3 的值。当其他条件不变时,b_3 与 0 的差别越大,说明交互作用越强。最流行的标准化测度是当其组成部分保持不变时,乘积项的半偏相关系数平方。这个值反映了因变量方差的多大比例可由乘积项单独解释,它可通过比较只含主效应模型的复相关系数平方和交互模型的复相关系数平方得到。在避孕措施的例子中,两项加叠模型(只含主效应模型)的复相关系数平方是 0.811,而三项交互模型得到的复相关系数平方是 0.937。因此,交互作用的“解释力”就是 $0.937-0.811=0.126$。这一交互作用在样本中可解释采取避孕措施的意愿的 12.6%的方差。这一指标

正向有偏，但当样本规模增大以及 R^2 更高时，这一偏差会相应降低。

对标准化作用大小测度的潜在缺点的讨论，请参见 McClelland 和 Judd(1993)以及 Jaccard(1998)的著作。

数字示例

一个具体的例子可以帮助我们总结前面的讨论。一项社会学研究用一个尺度来测量 100 名来自中西部社区天主教徒拥有大家庭的意愿(Y)。取值范围是 0 到 15，数字越大代表拥有大家庭的意愿更强。另外，受访者被问及在他们的成长过程中，家庭中孩子的个数，以及他们当前的家庭收入，收入以 1000 美元为单位(即 15 代表 15000 美元)。这三个测度的均值和标准差列在表 2.3 中。

表 2.3 变量均值与标准差

变量名	均值	标准差
理想的家庭规模	4.440	2.748
成长过程中的家庭规模	2.960	1.601
家庭收入	34.933	14.220

研究者假设那些成长于大家庭的个人相比那些成长于小家庭的个人，会对拥有大家庭持更正面的态度，因此就有更强的拥有大家庭的意愿。她还假设，这种成长过程中的家庭规模对理想的家庭规模的影响受到家庭收入的调节作用。其逻辑是，个体相对贫穷时，抚养小孩的成本就高，他们就不愿把对大家庭的渴望转化为实际。对富裕家庭而言，这种成本不构成限制条件。因此，研究者希望得到的结果是相比贫

穷家庭，在富裕家庭中，成长过程中的家庭规模对拥有大家庭意愿的作用更强。在此研究中，结果变量(Y)是拥有大家庭的意愿，而某人成长时的家庭规模是关键自变量(X)，收入则是调节变量(Z)。

在此分析中，X 和 Z 的值都进行了对中处理，以避免多重共线性问题，并使系数 b_1 和 b_2 更易理解。我们计算每一个个案对中后的 X 和 Z 的值，并使用 SPSS 软件进行多元回归分析，用 Y 对 X、Z 和 XZ 进行回归(注意，下面 X 和 Z 都指对中转换后的值)。乘积项(XZ)和其组成项(X 和 Z)的相关系数非常小(r 分别是-0.01 和 0.019)。

交互模型的复相关系数是 0.725，回归方程是：

$$Y = 4.4279 + 0.81324X + 0.0997Z + 0.0149XZ + e$$

b_1、b_2 和 b_3 的标准误分别是 0.098、0.045 和 0.007。

b_3 的 t 检验得到一个统计显著的结果($t = 2.187$, $p < 0.04$)，这表明存在交互作用。b_3 的 95% 的置信区间是 0.001到 0.029。半偏相关系数平方表明交互作用的解释力是 0.017，这是将只含主效应模型与交互模型的复相关系数平方相减得到的。对前者而言，复相关系数平方是 0.509；对后者而言，复相关系数平方是 0.526。那么，$0.526 - 0.509 = 0.017$，交互作用可解释理想家庭规模方差的 1.7%。

交互作用的本质通过 b_3 体现。b_3 表明根据收入的不同，成长过程中的家庭规模(RAISE)与理想的家庭规模(FSI)的关系如何变化。家庭收入每增加1000美元(这对应 Z 一个单位)，RAISE 对 FSI 的影响会增加 0.0149 个单位。

为更直观地展示调节变量如何影响 FSI 与 RAISE 的斜率

关系，我们在三个不同收入取值上（“低”收入值、“中”收入值和“高”收入值）分别计算用 RAISE 预测 FSI 的简单效应。我们用低于收入均值一个标准差作为“低”收入的标准，而“中”收入指的是收入均值，“高”收入指的是在收入均值一个标准差之上。收入标准差是 14.220，那么“低”收入的标准就是 $34.933-14.220=20.713$，或 20713 美元；“中”收入就是 34.933或 34933 美元；“高”收入就是 $34.933+14.220=49.153$，或 49153 美元。RAISE 对 FSI 影响的斜率正是当收入为“中等”时取得的，因为我们已对数据进行了处理，即对 X 和 Z 都进行了对中处理。b_1 是0.813，b_1 的标准误是0.098，t 值是 4.43($p<0.05$)。b_1 的 95%的置信区间是0.620到 1.009。

大家可利用前面讲解的方法分离“低”和“高”收入时相应的统计量，但我们在此使用的是变换策略。对“低”收入而言，用原收入分数减去 20.713，然后用变换后的分数乘以原先的 X 值，再用 Y 对这些相关的解释变量（X、Z_t 以及 XZ_t）进行回归，最后看 b_1 的值。同样，重复这一步骤，计算“高”收入分数（$34.933+14.220=49.153$），然后将原收入分数减去 49.153。经过这些步骤得到 b_1 的分析结果如下：

收入水平	b_1	标准误	95%置信区间	t	p 值
低	0.602	0.136	0.333 到 0.873	4.43	<0.001
中	0.815	0.098	0.620 到 1.009	8.33	<0.001
高	1.026	0.136	0.757 到 1.297	7.55	<0.001

在“低”收入水平上，成长过程中的家庭每多一个孩子，可使其理想的家庭规模增大 0.602；在“中”收入水平上，成长过程中的家庭中每多一个孩子，可使其理想的家庭规模增大 0.815；在“高”收入水平上，成长过程中的家庭中每多一个孩

子，可使其理想的家庭规模增大 1.026。成长过程中的家庭规模对理想的家庭规模的影响统计显著，这一点可由 RAISE 作用随收入的增大而增大得到印证，也与研究者假设的 b_3 取值一致。这些简单效应提供给读者关于交互作用的直观感受，但它们不代表对交互作用的正式检验。交互作用的检验基于对 b_3 显著性的检验。

图表展示

有些研究者喜欢用图来展示交互作用。方法之一是根据 Z 值的"低"、"中"、"高"分别画三条 X 对 Y 作用的直线，如上面计算的那样。相应的斜率分别是 0.602、0.815 和 1.027。要画这种图，我们需要计算每一条回归线的截距。可以利用下面这个公式从最初的回归方程中计算得到：

$$\text{在 } Z \text{ 值上}, X \text{ 对 } Y \text{ 作用的截距} = a + b_2 Z \quad [2.11]$$

a 指的是交互模型的截距，b_2 是方程中与调节变量相关的回归系数。

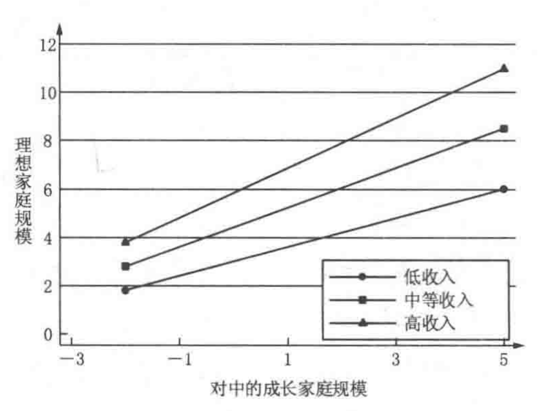


图 2.1　三个收入水平上成长过程中的家庭规模与理想的家庭规模的回归线

因为我们使用对中的 X 和 Z 值，因此，"低"收入 Z 是－14.22，"中"收入是 0，而"高"收入是 14.22。然后，把相应的数字代入方程 2.11，得到：

$$-14.22\text{ 时的截距} = 4.4279 + 0.0997(-14.22) = 3.0102$$

$$0\text{ 时的截距} = 4.4279 + 0.0997(0) = 4.4279$$

$$-14.22\text{ 时的截距} = 4.4279 + 0.0997(14.22) = 5.8456$$

图 2.1 中的回归线确认了三个方程 $Y = 3.0102 + 0.603X$、$Y = 4.4279 + 0.815X$ 和 $Y = 5.8456 + 1.026X$ 分别对应"低"、"中"、"高"三个收入 Z 值。记住，X 在这些方程中都指对中后的 X 值。若无交互作用，那么这三条线应当是平行的。很明显，事实并非如此。

第 3 节 | 一个定性解释变量与一个连续解释变量

为了展示一个定性解释变量和一个连续解释变量的情况，我们虚构一个样本。研究者关心青少年对自己与母亲关系的满意程度、该满意程度与性别以及与母亲共处时间之间的关系。

结果测度，即青少年的满意度，是一个从 0 到 21 的测度，用于测量他们对与母亲的关系的满意程度。分数越高，表明满意程度越高。青少年的性别是虚拟变量，1 = 男性，0 = 女性。母亲与孩子的共处时间是一个母亲自我报告的测度，母亲报告通常一个星期她与孩子共处的小时数。共处时间的均值是 24.06 小时，标准差是 4.62，样本量是 100。首先，我们需要考虑是否可以这样概念化：把共处时间作为关键自变量，把性别作为调节变量，然后修改这两个变量的角色，再重复分析一次。

定性调节变量

在此框架下，我们感兴趣的是青少年与母亲的共处时间对男性和女性青少年对其与母亲关系满意度的不同影响。

更正式的表述是，当用青少年满意度对共处时间进行回归时，共处时间的系数在男性和女性之间是否存在差异？如果两个斜率相同，那么共处时间对青少年满意度的影响对男性和女性来说就是一样的，并不存在交互作用。然而，若斜率不同，那么性别就调节了共处时间对青少年关系满意度的影响。

首先，对中时间解释变量，然后，满意度评分(Y)对对中后的共处时间(X)、性别虚拟变量(Z)以及乘积项(XZ)进行回归。请注意，虚拟变量并未对中。复相关系数的平方为0.784，且统计显著，即 $F(3, 97) = 116.07$, $p < 0.01$。回归方程是：

$$Y = 7.645 + 0.247 Time_c + (-5.270) Gender + (-0.260)(Time_c)(Gender) + e$$

交互作用的回归系数是−0.260，其标准误是0.113，95%的置信区间是−0.484到−0.036。t 检验是统计显著的($t = 2.30$, $p < 0.05$)，表明存在交互作用。乘积项的半偏相关系数平方是0.012，表明交互作用项可解释结果变量样本方差的1.2%。接下来，我们简短地考虑一下对 b_3 的解释。

解释变量 $Time_c$ 的回归系数0.247指的是当 $Gender = 0$ 的时候，共处时间对青少年满意度的作用。因为性别为0对应女性，那么这一系数就是共处时间对女性青少年满意度的影响：与母亲的共处时间每增加一小时，青少年满意度就会增加0.247个单位。这一系数的标准误是0.070(电脑输出结果)，95%的置信区间是0.108到0.386，t 检验统计显

著（$t = 3.53$，$p < 0.01$）。

上面得到女性共处时间对满意度的影响，而我们还想知道男性的情况。一个简单的方法是对虚拟变量重新编码，使女性=1，男性=0，然后用新的变量值乘以对中的时间变量，再重新计算。回归方程变成：

$$Y = 2.375 + (-0.013)Time_c + (5.270)Gender_r + (0.260)(Time_c)(Gender_r) + e$$

请注意，b_3 的绝对值没变，但正负号反转了。这就是逆向编码的结果，原因在下文会提及。共处时间系数为 -0.013，这是性别变量逆向编码后，共处时间对青少年满意度的影响。这里，性别为 0 代表男性，因此是共处时间对男性满意度的影响：与母亲的共处时间每增加一小时，会使男孩的满意度轻微变动，即减少0.013个单位。这一系数的标准误是 0.089（电脑输出结果），95%的置信区间为 0.189 到0.163，t 检验的结果不显著（$t = 0.15$，$p > 0.88$）。

这一关于简单效应的例子表明，青少年满意度对男性共处时间回归的结果是 -0.013，而对女性共处时间回归的结果是 0.247。若计算这两个斜率的差异，得到 $-0.013 - 0.247 = -0.260$。请注意，这正是最初回归分析中乘积项的系数，即 $b_3 = -0.260$。在传统的乘积项分析中，当乘积项涉及一个连续变量（X）和一个虚拟变量（Z）时，乘积项的系数总是等于 Y 对 X 进行两次回归、虚拟变量赋值为 1 的组的（第一次回归）系数减去原参照组重新编码被赋值为 1 的（第二次回归）系数。对 b_3 的显著性检验代表了对上述系数差异

的检验，这是单自由度的交互比较。在第一组分析中，b_3 对应 $-0.013-0.247=-0.260$；在第二组分析中，性别被重新编码，b_3 对应 $0.247-(-0.013)=0.260$。因此，b_3 的正负号发生变化。

总而言之，使用乘积项回归来计算女性 Y 对 X 回归的斜率，然后对虚拟变量重新编码，再计算男性 Y 对 X 回归的斜率。我们刻画了每一个斜率的显著性检验和置信区间，同时通过乘积项的回归系数(b_3)，我们正式检验了斜率的差异。

我们可以利用调节变量分组，在同一张图中画出不同的回归线。图 2.2 正是这么做的。交互作用表示不平行的两条直线。给定某组回归线的截距，就是当该组为参照组时，整体回归方程的截距。

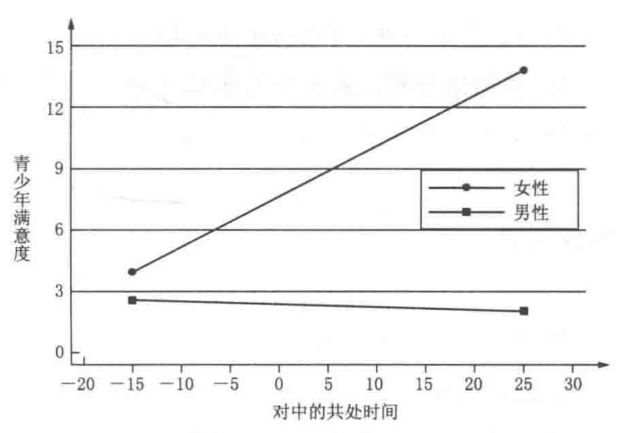


图 2.2 女性和男性共处时间对关系满意度的影响回归线

对女性而言，它是 7.645；对男性而言，它是 2.375(请注意，在图 2.2 中，$X=0$ 时的截距出现在图的中央，因为出于

美观考虑，我们把 Y 轴左移了）。

在应用研究中，科学家有时会单独计算每一组的回归方程并分别检验 Y 对 X 回归的斜率和显著性，以进行组间比较。当斜率在一组中统计显著，而在另一组中不显著时，就存在组间差异。这种分析往往有很多问题。重要的是，这种分析不能对两组之间的斜率差异进行正式检验，而如果我们要比较组间差异，这种检验就是必须的。一个斜率在一组（的回归）中统计显著，而在另一组（的回归）中统计不显著，但当两组进行正式比较时，斜率差异依然可能不显著。有可能两个斜率都统计显著，但正式的斜率差异统计检验也可能不显著，或者两个斜率都统计不显著，但两者的差异却是统计显著的。研究者需要进行正式的斜率差异比较才能谈论斜率的组间差异，而乘积项策略正是这么做的。另外，分组比较策略的问题是，当进行统计显著性检验时，它忽略了有用的残差信息。在乘积项分析中，给定解释变量上的残差是合并分析的，是根据分析涉及的所有组的信息而给出的。在分组分析中，方差只基于单个组。

连续调节变量

假设研究者进行不同的概念化，并调换关键自变量和调节变量的角色。现在考虑的是性别差异对青少年满意度的影响，且这些差异是否为关于与母亲共处时间的一个函数。这一分析使用同一个回归方程，即青少年满意度对对中后的时间变量、性别以及两者乘积项进行回归。我们关心使用不同的概念化和刻画方式时，模型体现出的不同特征。回想一

下，回归方程是：

$$Y = 7.645 + 0.247 Time_c + (-5.270) Gender + (-0.260)(Time_c)(Gender) + e$$

与前面一样，b_3 包含交互作用信息，即性别差异如何作为共处时间的一个函数而变动。接下来，我们来看此种理解方式。首先，检查性别的系数。因为性别是一虚拟变量，它的系数代表一个平均的差异，即它是 $Time_c = 0$ 时，男性（虚拟变量中编码为 1 的组）的平均满意度与女性的平均满意度之差。由于 $Time_c$ 是对中后的变量，那么 $Time_c$ 为 0 就对应了原共处时间变量的样本均值（即 24.06 小时）。男性和女性的平均满意度差异（－5.270）是当共处时间为其均值（24.06）时得到的。根据对 b_2 的 t 检验，这一性别差异统计显著（$t = 10.28$，$p < 0.01$），而这一差异的标准误是0.513，95%的置信区间是－6.288 到－4.253。截距项是 7.645，这是当所有解释变量都为 0 时，Y 的平均值。如果 $Time_c$ 为 0，性别为 0，那么方程中所有项都为 0，因此 7.645 就是当 *Time* 变量等于其样本均值时，女性的平均满意度。如果男性和女性的平均差异为－5.270，而女性的平均值是 7.645，那么简单的代数计算就可告诉我们男性的平均值。它应当是这一差异加上女性的均值，即 $-5.270 + 7.645 = 2.375$。注意，2.375（男性均值）减去 7.645（女性均值）等于平均差异，即－5.270。

若性别差异的估计值对满意度的影响在共处时间的不同取值下结果不同，又会如何？从上面的分析可知，共处时间为 24.06（小时）时，性别差异是 5.270。在其他取值下，情

况又如何？为便于理解，让我们计算共处时间为 25.06（小时）时（即共处时间增加一个单位时）的性别差异。我们利用变换策略。首先，把 *Time* 变量进行变换，使 0 值对应原时间变量的 25.06，即将原时间变量减去 25.06。变换后的时间变量为 $Time_t$，然后乘以虚拟变量性别，这几项同时进入回归方程，得到：

$$Y = 7.892 + 0.247 Time_t + (-5.530) Gender + (-0.260)(Time_t)(Gender) + e$$

性别系数变成－5.530，反映的是当 $Time_t$ 为 0 时，平均的性别差异。这意味着，当 $Time = 25.06$（小时）时，性别差异是－5.530。我们将之与先前的分析（共处时间为 24.06 小时，性别差异是－5.270）比较。增加一个时间单位到 25.06小时，这一差异变化了－0.260 个单位，变成了－5.530。现在再看 b_3。b_3 等于－0.260，这正是改变调节变量一个单位所带来的平均差异的变化。在传统乘积项分析中，当乘积项涉及一个连续变量（Z）和一个虚拟变量（X）时，给定连续变量增加一个单位，乘积项系数反映的是在虚拟变量赋值为 1 的组减去参照组的平均差异的变化。第二项分析中的截距项 7.892 是当 $Time = 25.06$ 时，女性的平均满意度。同样，可利用代数计算此例中男性的平均满意度，它就是平均差异加上女性的平均满意度，即 $-5.531 + 7.892 = 2.361$。

尽管回归方程包含了所有的相关信息，有些研究者报告的不仅是回归方程及其相关统计量，而且还报告了在两个或三个给定连续性调节变量取值上，关键自变量的平均差异。

这有助于为读者提供直观的对交互作用的感受。例如，我们使用变换策略得到下面这个统计表：

共处时间	男孩均值	女孩均值	平均差异	95%置信区间	t	p值
20小时/周	2.43	6.64	−4.21	−5.54到−2.98	4.43	<0.001
25小时/周	2.36	7.88	−5.52	−6.59到−4.44	8.33	<0.001
30小时/周	2.30	9.11	−6.82	−8.63到−5.00	7.55	<0.001

当每周共处时间为20小时时，性别差异对关系满意度的影响是−4.21(即女性比男性的满意度高4.21个单位)；当每周共处时间增加到25个小时时，性别差异变成−5.52；若每周共处30个小时，差异增加更多，达到−6.82。这一展示性统计提供了直观的交互作用效应，我们看到，均值差异随着调节变量取值的不同而变动。然而，应当谨记，这些简单效应不是交互作用检验的核心。交互作用的正式检验应当在b_3中。

这一结果可用直方图展示，如图2.3。若不存在交互作用，那么每一个共处时间点上男性和女性柱高度的差异就

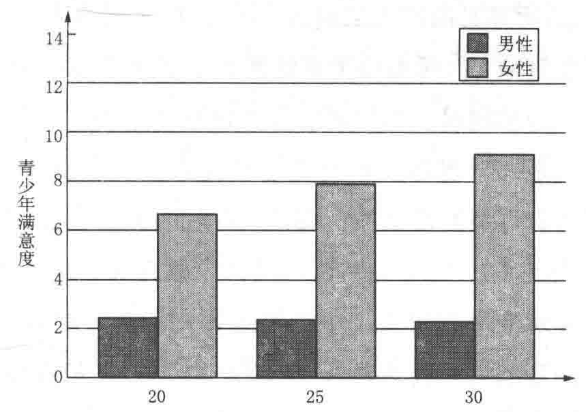


图2.3 选定共处时间点上不同性别间关系满意度差异的均值

应相同。缺乏这种一致性就表明存在交互作用。该图表明,当与孩子共处的时间增加时,性别差异对关系满意度的影响变大了。另外一个图形展示方法是画出由关键自变量定义的不同组之间,Y 对 Z 的单独的回归线。这已由图 2.2 进行了展示。两条回归线之间的距离在水平线(调节变量)上的差异反映了 Y 在这一点上的平均差异。图 2.2 包含的信息比图 2.3 更多,但也有些研究者认为图 2.3 更有意思。

两组以上的定性变量

有一些应用涉及定性和连续性解释变量,且定性变量包含两个以上的层级。上面这些一般化的原则都适用于这种情况。

请考虑这样的情况:结果变量是收入(Y),两个解释变量是受教育年限和种族。假设种族有三类:美国黑人、美国白人[8]、拉美裔。这由两个虚拟变量表示。假设第一个虚拟变量是 D_{AA},所有美国黑人赋值为 1,而其他人赋值为 0。第二个虚拟变量是 D_L,所有拉美裔赋值为 1,其他人赋值为 0。以美国白人为参照组。为了表示两个解释变量的交互作用,我们构建代表一个解释因素的变量和代表另一个解释因素变量之间所有可能的乘积项。假设 ED 代表教育测度,我们就得到乘积项$(ED)(D_{AA})$和$(ED)(D_L)$。教育和种族整体的交互作用要求使用方程 1.1 中的层级回归检验。只含主效应模型中的复相关系数平方(只含 ED、D_{AA}和 D_L 的模型)与包含主效应和所有乘积项模型的复相关系数平方进行比较。

仅根据乘积项单个 b 系数的显著性检验来检查整体交互作用的显著性是不够的，因为整体交互作用具有一个以上的自由度，必须使用层级策略检验。

研究者如何理解回归方程取决于挑选哪个变量作为调节变量这一问题？假设在第一个例子中，种族是调节变量（调节变量是定性变量）。在此情况下，研究者感兴趣的是教育对收入的作用，且这种作用如何随种族变化而变化。回归方程的一般形式是：

$$Y = a + b_1 ED + b_2 D_{AA} + b_3 D_L + b_4 ED(D_{AA}) + b_5 ED(D_L) + e \quad [2.12]$$

正如前面所讨论的，乘积项回归系数反映了斜率的差异。b_4 关注 Y 对 ED 回归的斜率，并估计了 D_{AA} 为 1 的组的斜率减去参照组斜率的差异。换言之，它反映了美国黑人与美国白人的斜率差异。b_5 对应同样的斜率差异，指拉美裔和美国白人的斜率差异。显著性检验、标准误、置信区间提供单自由度交互比较通常所关心的额外信息。

那么，如果除了由 b_4 和 b_5 反映的这两个比较外，我们还想评估 Y 对 ED 回归时，美国黑人和拉美裔的斜率差异呢？这种比较和它的相关显著性检验可以通过产生一组新的虚拟变量和乘积项，或把参照组改成美国黑人或拉美裔，然后在新回归方程中检查乘积项系数，就能分离出来了。第 4 章会讨论如何进行这种多元比较，并调整多重比较带来的扩大的误差率。

在方程 2.12 中，ED 的系数 b_1 是一个简单效应，它指的是当所有种族虚拟变量都为 0 时，教育对收入的作用。

$D_{AA}=0$ 且 $D_L=0$ 指参照组，因此 b_1 是在美国白人中，教育对收入的作用。标准误、置信区间和显著性检验都针对这一作用，可从 b_1 的电脑输出结果中看到。如果需要分离出拉美裔相应的简单效应，那就需要重新定义一组新的虚拟变量和乘积项，使得拉美裔成为参照组，再重新进行电脑分析。此过程与美国黑人的分析过程相似。

请考虑这样的情况：教育是调节变量。现在我们感兴趣的是种族差异对收入的影响，且这些差异如何受到教育的影响。使用同样的回归方程，但研究者在使用这种方法之前，通常会对中教育（或者把教育对中在一个有意义的值上）。下面讨论中的对中操作均如此。

首先，我们考虑 D_{AA} 的系数，即 b_2。这一系数是当 $ED_c=0$ 时，D_{AA} 这一虚拟变量取值为 1 的组和参照组的平均收入差异，这个差异就是当教育等于样本均值时，美国白人和美国黑人之间平均的收入差异。D_{AA} 和 ED_c 的乘积项系数 b_4 反映了教育增加一个单位所导致的平均收入差异的变化。D_L 的系数是 b_3，代表当教育等于样本均值时，拉美裔和美国白人直接的平均收入差异。D_L 和 ED_c 的乘积项系数 b_5 反映了教育增加一个单位时，拉美裔和白人的平均收入差异的变动情况。如果研究者对拉美裔和美国黑人可比的系数和显著性检验感兴趣，那么只需重新定义虚拟变量，使美国黑人或拉美裔成为参照组，然后再进行电脑分析。使用上述讨论的原则就可分离出相应的系数。

回归方程中的截距是当 $ED_c=0$ 时，参照组（美国白人）的平均收入。

交互的形式

表 2.4　两种交互作用

Z 值	例 A:两组之间平均差异	例 B:两组之间平均差异
1	2.0	2.0
2	4.0	4.0
3	6.0	6.0
4	8.0	8.0
5	10.0	10.0
6	12.0	10.0
7	14.0	10.0

如前所述,当乘积项中两个变量都是连续变量时,传统交互模型评估的是一种特殊的交互形式,即双线性交互。当乘积项中的一个是定性变量,另一个是连续变量时,传统交互模型估计的是一种特定的交互。这个模型可通过两种不同的情况来说明,尽管两种情况反映了同一个模型。一种情况是调节变量为定性变量。在这种情况下,预先假定在调节变量的每个类别上,Y 和 X(结果变量和连续性关键自变量)的直接关系为线性关系都为真。如果至少在调节变量的某一个类别上,Y 和 X 的关系是非线性的,那么传统乘积项模型的设定就是错的,此时就需要另一模型(见第 4 章的讨论)。另一种情况是调节变量为连续变量。在此情况下,预先假定组均值的有序变化是 Z 的线性函数。请考虑表2.4中的两个例子。例 A 表示的是两组之间的差异是 Z 的线性函数:Z 每增加一个单位,组间的平均差异就增加两个单位。而例 B 则反映了非线性关系不能用传统的乘积项方法来建模。在 Z 的低层级上,Z 每增加一个单位,组间的平均差异就增加两个单位,但这最终渐近并平稳下来(不再增加)。

第4节 | 小结

研究者经常关注双向交互的交互模型。交互中所涉及的两个变量本质上都可以是连续的,或一个连续、一个定性。在这两种情况下,研究者都宣称一个是关键变量,另一个是调节变量。恰当的乘积项可以通过计算得到,然后交互模型就包括了"主效应"和乘积项。当整体交互作用是单自由度时,整体显著性检验就反映在乘积项的回归系数上。当整体交互有一个以上的自由度时,就需要运用层级检验比较只含有主效应模型和交互模型的差异来估计整体交互作用的统计显著性。

非标准化系数的交互效应的优势在于,其回归系数反映了单自由度交互作用。标准化回归系数的优势在于,交互项的半偏相关系数平方反映了交互作用的大小。

典型的对交互作用的理解关注的是与关键自变量相关的回归系数以及与乘积项相联系的回归系数,这取决于哪个变量是调节变量。不同的对中策略和重新编码方法可帮助读者理解交互作用的本质。

第3章

三向交互作用

本章把第 2 章的原则扩展到三向交互作用分析上。首先考察全部是连续变量的情况，然后考察定性和连续变量混合的情况。

第1节 | 连续解释变量

假定有一项研究,一名行为医学研究者关注父母为了预防某种形式的肝炎,给孩子注射疫苗的意愿。这种意愿测度的取值是0到100,分数越高说明接种疫苗的意愿越强烈。研究者研究三个影响接种疫苗意愿的潜在变量。第一个变量是如果父母不给孩子接种疫苗,预计孩子感染这种疾病的可能性。我们称这个变量为疾病的"预期易染性",用0到100的分数进行测量,分数越高,疾病传染性越强。第二个变量是如果孩子真的感染了某种肝炎,那么父母预计这种病会有多严重的后果。这一变量被称为"预期严重性",也用0到100分测量,分数越高,说明预期的严重性越强。第三个变量是预计接种疫苗使孩子避免染病的可能性。这个变量又被称为"预期有效性",用0到100分测量,分数越高,说明预期有效性越强。接种疫苗的意愿受到这三个变量交互函数的影响。此研究的样本是200名家长。

为分析三向交互作用,我们需要确定一个关键自变量和两个调节自变量来明确交互的逻辑。对三向交互作用而言,需要区分两个调节变量。假设研究者选择的关键自变量是预期易染性,并假设接种疫苗的意愿是预期易染性的函数。具体而言,假设预期易染性越高,接种的意愿也会越强。进

一步的假设是，预期易染性对接种意愿的作用会受到预期严重性的调节作用。如果父母认为这个疾病并不严重，那孩子易受传染这一问题就无关紧要了。父母缺乏动力去为他们认为无关痛痒的疾病寻求疫苗。然而，当预期严重性增强时，预期易染性就会影响接种疫苗的意愿。在此情况下，预期严重性被称为“一阶调节变量”，因为它直接调节了预期易染性对接种疫苗意愿的影响。研究者进一步假设，预期严重性的调节作用会受到预期有效性的影响。如果父母认为疫苗不那么有效，那么无论多么容易感染这种疾病或者万一染病，后果有多么严重，他们都不会要求接种。然而，当疫苗的预期有效性增强时，预期严重性和预期易染性的交互机制会自我增强。疫苗的预期有效性是二阶调节变量，因为它调节了一阶调节变量对关键自变量和因变量的作用。当然，这并非三向交互作用必须的概念化方式。然而，我们发现，一阶和二阶调节变量的理解方式对思考三向交互作用很有用。并且，当调查者描述三向交互作用时，他们总是采用这种方式（尽管有时候不那么明显），从而使复杂的关系变得容易理解。

为了分析传统的交互模型，需要形成乘积项并加入主效应模型反映双向交互和三向交互。主效应模型是：

$$Y = a + b_1X + b_2Z + b_3Q + e$$

在这里，X 是预期易染性，Z 是预期严重性，Q 是预期有效性。三向交互作用模型加入了所有这三个解释变量成对的项（XZ、XQ 和 ZQ）以及三个解释变量共同的乘积项（XQZ）。模型就变成如下形式：

$$Y = a + b_1 X + b_2 Z + b_3 Q + b_4 XZ + b_5 XQ + b_6 QZ + b_7 XZQ + e \quad [3.1]$$

三向交互作用的显著性就是 b_7 的显著性。用标准术语来讲，三向交互作用的解释力是用方程 3.1 的复相关系数平方减去一个基于方程 3.1 但省略 XZQ 这一项的方程的复相关系数平方。低阶系数的理解总是条件性地取决于高阶乘积项，而条件化是指其他变量在高阶交互项上等于 0。例如，b_4 代表当 $Q = 0$ 时，X 和 Z 对 Y 的双向交互作用；b_1 代表当 $Z = 0$ 且 $Q = 0$ 时，X 对 Y 的作用。关键自变量、一阶调节变量、二阶调节变量的区分有助于我们理解方程的系数。

回到例子中，先对中所有解释变量，然后再形成相关乘积项。SPSS 回归输出结果列在表 3.1 中。三向乘积项的系数是 0.0008137，且统计显著（$t = 4.715$，$p < 0.01$，95% 的置信区间是 0.000473 到 0.001154），表明存在三向交互作用。而省略三向乘积项方程的复相关系数平方是 0.400，完整模型的复相关系数平方是 0.462。那么这两个复相关系数平方的差异为 0.062，表明三向交互作用能够解释接种疫苗意愿的方差的 6.2%。

基于前面展示的原则，预期易染性（X）系数是 0.219，这代表预期严重性（Z）和预期有效性（Q）都为 0（即当预期严重性和预期有效性为“平均”或“中等”，因为对它们进行过对中处理）时，预期易染性对接种意愿的影响。0.219 表明预期易染性每增加一个单位，接种意愿就增加 0.219 个单位（$t = 5.062$，$p < 0.01$，95% 的置信区间是 0.134 到 0.304）。而 XZ 的系数反映了当二阶调节变量预期有效性为 0（即预期有效性是“中等”或“平均”）时，预期易染性与一阶调节变量

表 3.1 SPSS 三向交互作用输出结果

模 型	R	R^2	调节的 R^2	测量的 SE
1	0.680	0.462	0.443	9.85080

模 型	非标准化回归系数		标准化回归系数	t	Sig.	B 的 95%置信区间	
	B	SE	β			低 端	高 端
(常数项)	33.332	0.697		47.849	0.000	31.958	34.706
易染性(X)	0.219	0.043	0.270	5.062	0.000	0.134	0.304
严重性(Z)	0.203	0.044	0.245	4.629	0.000	0.116	0.289
有效性(Q)	0.248	0.044	0.299	5.635	0.000	0.161	0.334
XZ	0.01227	0.003	0.246	4.606	0.000	0.007	0.018
XQ	0.01254	0.003	0.243	4.519	0.000	0.007	0.018
QZ	0.01383	0.003	0.263	4.951	0.000	0.008	0.019
XQZ	0.0008137	0.000	0.254	4.715	0.000	0.000	0.001

预期有效性的交互作用。这一系数是 0.01227（$t=4.61$，$p<0.01$，95% 的置信区间是 0.007 到 0.018），这是预期有效性为“平均值”时，预期易染性每增加一个单位，预期易染性对接种意愿斜率的增加值。控制预期有效性为其样本均值时，预期严重性每增加一个单位，预期易染性对接种意愿的影响力就增加 0.01227 个单位。这一系数的理解与第 2 章中的双向交互作用一样，但它是以 $Q=0$ 为条件的（在此例中，代表对中后的预期有效性为 0）。

为理解三向乘积项系数的意义，我们要重新计算当预期有效性（Q）比其均值高一个单位而非其均值时，此例中 XZ 的系数。将预期有效性减去一个常数 45.79，而非减去样本均值 44.79 即可。这样，我们把变换后的预期有效性 0 值定义为 45.79，而非原尺度中的 44.79。然后，使用变换后的分数重新计算乘积项，并重新进行回归。

正如第 2 章讨论的，复相关系数平方和 b_7 的值在这个变换之后保持不变。在新分析中，XZ 乘积项系数为 0.0130837。通过计算前后两次回归中 b_4 的差异来看双向交互系数变化，即 $0.0130837-0.01227=0.0008137$。检查表 3.1中的三向交互作用系数，正好是 0.0008137，正是预期有效性变动一个单位时，双向交互作用变动的量。对一个由三个连续变量组成的三向交互模型而言，关于 X、Z 和 Q 以及它们之间不同的乘积项，我们令 X 是关键自变量，Z 是一阶调节变量，而 Q 是二阶调节变量。三向交互作用的系数是给定 Q 变动一个单位，X 和 Z 双向交互作用系数的变动数量。如果重新回归，使预期有效性（Q）高于其样本均值两个单位进行对中处理，那么 XZ 的系数就变成：

$$0.01227 + 0.0008137 + 0.0008137 = 0.0138974$$

尽管回归方程中包含了所有这些理解三向交互作用所需的相关信息，但大部分读者还是感到从方程本身获取三向交互作用的直观理解依然有些困难。正如在第 2 章中我们发现，提供不同的 Z 和 Q 取值组合时，X 对 Y 作用的斜率有助于读者思考交互作用。其结果列在表 3.2 中。此表有四种情况，在这些情况下观察 Y 对 X 回归的斜率：(1) Z 值低，Q 值低；(2) Z 值低，Q 值高；(3) Z 值高，Q 值低；(4) Z 值高，Q 值高。低和高都被定义为低于和高于 Q 和 Z 各自样本均值的一个标准差。Y 对 X 在每种情况下的回归斜率展示在一个 2×2 的析因分析表中，行的值代表一阶调节变量的低和高，而列的值代表二阶调节变量的低和高。这些系数的值的显著性检验及其置信区间都通过计算机计算得到，并通过变换策略分离出相关系数，观察其对预期易染性的影响。此表说明，预期有效性相对较低时，预期易染性对接种疫苗意愿作用的影响可忽略不计，且相比预期有效性相对高时，意愿要小得多。事实上，只有当预期严重性和预期有效性都很高时，才能看到预期易染性对接种意愿有实际意义的影响力。

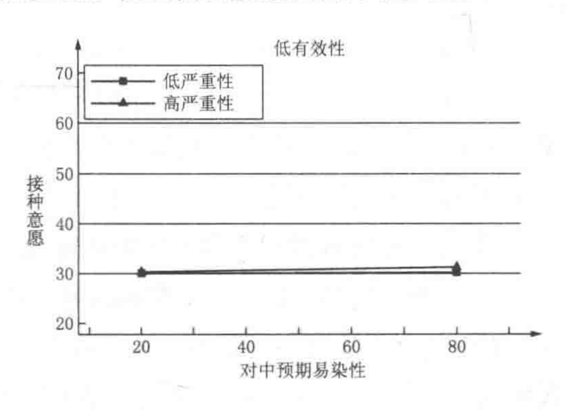
表 3.2 作为预期严重性和预期有效性函数的接种意愿对预期易染性回归的斜率

	低有效性			高有效性		
	斜率	95%置信区间	t	斜率	95%置信区间	t
低严重性	0.003	−0.128 到−0.188	0.376	0.009	−0.183 到 0.200	0.089
高严重性	0.016	−0.145 到 0.178	0.197	0.820	0.648 到 0.993	9.39*

注：* 表示 $p<0.01$。

这个统计结果可用一个并排图形来展示。一张图是当

Q为"低"时，通过Z的两个值得到的回归线；另一张图是Q为高时，通过Z的两个值得到的回归线。这两张图就可并排展示。图3.1正是采用这种展示方法。若不存在三向交互作用，那么这两个斜率在一张图上将是平行的，而在另一张图上也应如此。但我们的数据表明并非如此。



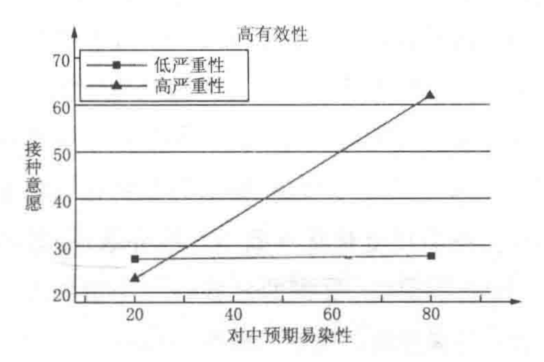


图3.1　作为预期严重性和预期有效性函数的预期易染性对接种意愿的作用的回归线

为构建此图，我们需要计算给定的Z值和Q值时，Y对X回归的截距。截距就是在完整的交互回归中，使用变换后的Z值和Q值分离出感兴趣斜率的截距项。

第 2 节 | 定性和连续解释变量

假设一位研究者对司法系统中的偏见感兴趣。他设计了一个研究，使参与者阅读一段针对被告所作出的法庭案例分析报告，然后给出此被告被判定为有罪的可能性分数，分数由 0 到 100。结果变量从 0 到 100，分数越高表示越高的有罪概率或越高的被判定有罪的可能性。所有参与者都阅读相同的报告，只不过其中一半人被告知被告是美国黑人，而另一半人则被告知被告是美国白人。参与者中有一半是美国黑人，而另一半是美国白人。这形成一个 2×2 的析因设计，交叉的是被告的种族和"陪审员"(研究参与者)的种族。研究者同时测量了第三个变量，即参与者具有自由主义或保守主义倾向，相关尺度是从 0 到 100 的分数，分数越高表示越强的自由主义倾向。被告的种族由一个虚拟变量 AC 表示，美国白人是参照组。陪审员种族同样由一个虚拟变量表示——JU，美国白人是参照组。自由主义倾向尺度 L 进行对中操作，由此产生相应的乘积项(因此，此处 L 代表对中后的自由主义—保守主义尺度)。表 3.3 是 SPSS 回归程序的输出结果。

三向乘积项的系数 1.007 统计显著($t = 7.31$, $p < 0.01$, 95% 的置信区间为 0.735 到 1.279)，表明存在三向交

互作用。省略 $AC\times JU\times L$ 模型的复相关系数平方是0.595，全模型的复相关系数平方是0.782。两者之差为0.192，表明三向交互作用可以解释19.2%的定罪方差。接下来，我们在两种情况下讨论这些系数：一是关键自变量为连续变量；二是关键自变量为定性变量。

连续关键自变量

在此框架下，研究者感兴趣的是自由主义—保守主义倾向对定罪的影响。研究者指出，更具自由主义倾向的陪审员相比保守的陪审员，较不倾向于判定被告有罪。然而，这种自由主义—保守主义倾向对定罪的影响会受到被告种族的调节作用。当被告是美国黑人时，相比被告是美国白人，自由主义—保守主义倾向将会是一个对定罪更有力的解释变量。因此，被告种族就是一阶调节变量。根据三向交互作用原理，被告种族对自由主义—保守主义倾向对定罪的调节作用还取决于陪审员自身的种族。研究者预测，前述(被告)种族和(陪审员)自由主义—保守主义分数的双向交互作用的机制在美国白人陪审员中(相比美国黑人陪审员)更为显著。因为在司法案件中，相比美国黑人陪审员，偏见和种族被认为在美国白人陪审员身上体现得更为明显。陪审员种族因此就是一个二阶调节变量。

自由主义—保守主义系数在表3.3中是−0.576，反映的是当 $AC=0$ 且 $JU=0$ 时，自由主义—保守主义倾向对定罪的影响。这表示当美国白人被告受到美国白人陪审员审判时的情况，它是简单效应。如果先分离出其他三种情况下

表 3.3 陪审员例子中的 SPSS 输出

模　型	R	R^2	调节的 R^2	测量的 SE
1	0.884	0.782	0.775	8.58238

模　型	非标准化回归系数		t	Sig.	B 的 95%置信区间	
	B	SE			低　端	高　端
（常数项）	39.006	1.111	35.102	0.000	36.817	41.196
被告种族(AC)	17.235	1.570	10.980	0.000	14.142	20.328
陪审员种族(JU)	1.228	1.570	0.782	0.435	−1.865	4.320
自由主义(L)	−0.576	0.068	−8.473	0.000	−0.710	−0.442
$AC*JU$	−17.241	2.218	−7.772	0.000	−21.611	−12.870
$AC*L$	−0.952	0.098	−9.717	0.000	−1.145	−0.759
$JU*L$	−0.033	0.098	−0.340	0.734	−0.226	0.160
$AC*JU*L$	1.007	0.138	7.307	0.000	0.735	1.279

的简单效应，可使方程中其他系数的含义更明确。这可通过为 AC 和 JU 重新计算一组新的虚拟变量而达成，即重新定义参照组，从而使与自由主义—保守主义相关的系数代表我们感兴趣的简单效应。例如，重新定义 AC 的参照组为美国白人，而 JU 的参照组为美国黑人，那么重新计算乘积项并重新进行回归，得到与 L 相连的系数就是美国白人被告被美国黑人评判时，自由主义—保守主义倾向对定罪的影响。表3.4展示了四种不同情况下的电脑分析结果。

表 3.4　自由主义对定罪的作用斜率

	斜率	标准误	95%置信区间	t	p 值
AA 陪审员，AA 被告	−0.554	0.066	−0.685 到−0.424	8.35	<0.001
AA 陪审员，EA 被告	−0.609	0.071	−0.748 到−0.470	8.64	<0.001
EA 陪审员，AA 被告	−1.528	0.071	−1.667 到−1.389	21.65	<0.001
EA 陪审员，EA 被告	−0.576	0.068	−0.710 到−0.442	8.473	<0.001

注：AA = 美国黑人；EA = 美国白人。

仔细检查表 3.3 中关键自变量与一阶调节变量乘积项（$AC \times L$）的系数。这一系数反映了一个定类变量和一个连续变量之间的双向交互作用，所以它反映了斜率差异。具体而言，它指 $JU = 0$，定罪分数对自由主义—保守主义倾向回归时，黑人被告与白人被告的自由主义变量的斜率差异。JU 为 0 代表陪审员为白人，因此只关注白人陪审员的情况。这一系数等于−0.952。查阅表 3.4 并只观察白人陪审员，我们发现自由主义—保守主义对黑人被告的归罪贡献系数是−1.525，而对白人被告是−0.576。两个斜率之差是 $(-1.528)-(-0.576)=-0.952$，这正是双向交互作用参数。

从表 3.3 可见，这一斜率差异统计显著（$t=9.72$，$p<0.01$，95% 的置信区间为−1.145 到−0.759）。对包含一个连续解释变量 X、两个定性解释变量 Z 和 Q 以及它们相应乘积项的交互模型来说，令 X 为关键自变量，Z 为一阶调节变量，而 Q 为二阶调节变量。对用虚拟编码方法的定性变量而言，乘积项 XZ 的回归系数就是斜率之间的差异。它关注虚拟变量 Z 赋值为 1 的组与参照组之间的 X 对 Y 作用斜率的差异，当然，这些都只针对 Q 为参照组的情况。

对白人陪审员而言，黑人被告和白人被告之间的斜率差异是−0.952，那么黑人陪审员的情况又如何呢？这个差异的值、标准误、置信区间都可通过重新定义 JU 这一虚拟变量，使美国黑人陪审员成为参照组而实现。再重新计算乘积项，重新进行回归，检查 $AC \times L$ 乘积项系数。在新的分析中，这个系数是0.055。这确认了表 3.4 中的相关的斜率是−0.554 和−0.609，他们的差异是 $(-0.554)-(-0.609)=0.055$。这个差异统计不显著（$t=0.565$，$p<0.57$，95% 的置信区间为−0.136 到 0.246）。

总的来说，对美国黑人陪审员而言，黑人被告 Y 对 X 的斜率减去白人被告 Y 对 X 的斜率，其差是 0.055，这可由美国黑人陪审员（与自由主义分数 L）的双向交互参数得到。美国白人陪审员的斜率差异是−0.952，它代表美国白人陪审员（与自由主义分数 L）的双向交互作用参数。如果不存在三向交互作用，那么这些双向交互参数应当相同（假设不存在样本偏差）。事实上，它们确实提示可能存在三向交互作用。[9]两个双向交互参数的差异是 $(0.055)-(-0.952)=1.007$。再看三向乘积项的系数，注意它等于 1.007。而这一

参数的显著性检验评估的是在假设双向交互作用无差异这一虚无假设为真时，观察到结果为1.007或者更大结果的概率。对有一个连续变量 X、两个定性变量 Z 和 Q 及其相应乘积项的交互模型，令 X 作为关键自变量，Z 为一阶调节变量，Q 为二阶调节变量。对虚拟编码的定性变量而言，三向乘积项的回归系数指的是斜率差异的不同。差异是当 Z 赋值为 1 的组的 Y 对 X 的斜率减去参照组的 Y 对 X 的斜率。这一斜率差异对虚拟变量 Q 而言，即当 Q 赋值为 1 的组的 Y 对 X 的斜率值减去 Q 为参照组的 Y 对 X 的斜率。

简言之，表 3.4 提供了不同的实验情境下 Y 对 X 的斜率，并确认了这些简单效应的显著性检验。三向交互作用隐含了双向交互作用会根据二阶调节变量取值的变化而变化。表 3.4 中斜率的模式很明显地展示了三向交互作用的来源。白人陪审员的斜率差异比黑人陪审员的斜率差异大得多。对应由 Z 定义的两组，在每个 Q 的取值上单独作图，进而得到的表示 Y 对 X 的回归线的并排图可展示这一点。

定性关键自变量

我们可使用定性解释变量作为关键自变量，重新分析上面的研究。概念化的逻辑是这样的：研究者预计由于负面的刻板印象和偏见，美国黑人被告会比美国白人被告更容易被判有罪。然而，这种倾向会受到陪审员种族的考验（一阶调节变量）。只有当陪审员是美国白人时，才会出现美国黑人被告比白人被告更易被判有罪的情况。这一假设源于白人陪审员会表示出偏见，而黑人陪审员则不会。最后，陪审员

种族的调节效应依赖于陪审员本身的自由主义—保守主义倾向。对那些自由主义的陪审员而言，这不是问题，即无论陪审员是黑人还是白人，都不会对黑人被告有偏见。

三向交互作用的检验与前文中的完全一样，我们这里仍用表3.3的结果。其焦点是使用上述概念框架，在不同的低阶系数上展示三向交互作用。

确认三向交互作用存在时，首先检查与关键自变量和一阶调节变量联系在一起的系数，即乘积项$AC \times JU$。这一系数是−17.241。如果乘积项中的两个自变量都是虚拟变量，那么这些相关系数可反映在一个2×2的表中。

表3.5 自由主义倾向分数取均值、定罪概率作为被告种族和陪审员种族的函数

	均值	标准误	95%置信区间
AA陪审员，AA被告	40.23	1.11	38.04到42.41
AA陪审员，EA被告	40.23	1.11	38.04到42.41
EA陪审员，AA被告	56.24	1.11	54.06到58.43
EA陪审员，EA被告	39.00	1.11	36.62到41.20

注：AA = 美国黑人；EA = 美国白人。

这个2×2表格交叉的是第一个虚拟变量取值为1时相比其参照组，以及第二个虚拟变量取值为1时相比其参照组。表3.5识别出这一子表。我们可识别每个单元格的平均被判有罪概率的均值，当然，条件是对中的自由主义—保守主义分数为0。回归全模型中的截距是当$AC=0$、$JU=0$且$L=0$时，截距的有罪判定的均值。因此，当自由主义—保守主义分数是平均值时，美国白人被告由美国白人陪审员来评分的平均有罪判定分数是39.006(95%的置信区间是36.817到41.196)。可通过代数(如第4章讨论的)或通过重

新编码虚拟变量并重新进行回归，从而得到其他三个单元格的均值。在后一个方法中，我们重新定义虚拟变量从而使参照组对应 2×2 表格中我们感兴趣的单元格，然后重新计算乘积项，重新进行回归，并注意新方程中截距项的值。使用这一策略，就会得到表 3.5 中四个单元格的均值及其置信区间。

如第 1 章所讨论的，2×2 表中交互作用的参数是调节变量在第一个类别上关键自变量取值的差异减去调节变量在第二个类别上关键自变量取值的差异，即 $(40.23-40.23)-(56.24-39.00)=-17.24$。请注意，这个值等于 $AC\times JU$ 的回归交互作用。标准误、置信区间和显著性检验都提供了检验 2×2 表格中交互作用有效性的视角。

对交互模型而言，有两个定性解释变量 X 和 Z、一个连续性解释变量 Q 以及它们之间产生的乘积项。令 X 为关键自变量，Z 为一阶调节变量，Q 为二阶调节变量。对虚拟编码的定性变量而言，XZ 的回归系数代表两个均值的差异。它是当 $Q=0$(对中后 Q)时，X 取值为 1 的组的均值减去参照组的均值之差，再减去 Z 取值为 1 的组的均值减去 Z 参照组的均值之差。

三向交互作用系数告诉我们给定的自由主义—保守主义分数变动一个单位，双向交互作用系数变动的数量。三向交互作用是 1.007，这表示自由主义倾向增加一个单位，前述交互作用参数 -17.241 将等于 $-17.241+1.007=-16.234$。如果自由主义倾向增加两个单位，那么前述交互作用参数将变成：

$$-17.241+1.007+1.007=-15.227$$

随着自由主义倾向的增强，被告种族和陪审员种族的交互作用的大小趋近于0。

正如前面那个例子，选定一些连续性二阶变量的值，展示相关的2×2表，这有助于读者理解三向交互作用。表3.6给出了自由主义—保守主义倾向的"低""中""高"值。"低"定义为比自由主义—保守主义分数均值低一个标准差，"中"是均值，"高"则是比均值高一个标准差。使用变换策略和"虚拟变量重新编码"策略来分离出相应的统计数值。如此安排，以便关键自变量由行来表示，而列是一阶调节变量，二

表3.6 被告种族与陪审员种族双向交互作用(AC×JU)对定罪概率影响的均值(在自由主义—保守主义倾向的不同层次上)

		AA陪审员	EA陪审员
低自由主义倾向	AA被告	49.18	80.92
	EA被告	50.07	48.31
		(49.18－50.07)－(80.92－48.31) =－33.50**	
		AA陪审员	EA陪审员
中自由主义倾向	AA被告	40.23	56.24
	EA被告	40.23	39
		(40.23－40.23)－(56.24－39.00) =－17.24	
		AA陪审员	EA陪审员
高自由主义倾向	AA被告	31.27	31.56
	EA被告	30.39	29.71
		(31.27－30.39)－(31.56－29.71) =－0.97**	

注：AA＝美国黑人；EA＝美国白人。

** 表示 $p<0.05$。

阶调节变量用于分隔这个 2×2 表。每一个 2×2 表中的单元格都是交互作用的参数。三向交互作用由此变得很明显，因为双向交互作用估计值在不同的自由主义—保守主义层次上变动。可用先前讨论的原则进行图表展示。

两个类别以上的定性变量

上述逻辑很容易扩展到有两个以上类别的定性变量的情况。我们必须谨记这些系数的条件性本质以及特定设计中单元格所代表的乘积项。例如，在第一个例子中，关键自变量是连续变量。假设被告的种族有三个类别（美国黑人、美国白人、拉美裔），并且陪审员也有三个类别（美国黑人、美国白人、拉美裔），那么全模型就是：

$$\begin{aligned} Y = a &+ b_1 L + b_2 AC_{AA} + b_3 AC_L + b_4 JU_{AA} + b_5 JU_L \\ &+ b_6(L)(AC_{AA}) + b_7(L)(AC_L) + b_8(L)(JU_{AA}) \\ &+ b_9(L)(JU_L) + b_{10}(AC_{AA})(JU_{AA}) + b_{11}(AC_{AA})(JU_L) \\ &+ b_{12}(AC_L)(JU_{AA}) + b_{13}(AC_L)(JU_L) \\ &+ b_{14}(L)(AC_{AA})(JU_{AA}) + b_{15}(L)(AC_{AA})(JU_L) \\ &+ b_{16}(L)(AC_L)(JU_{AA}) + b_{17}(L)(AC_L)(JU_L) + e \end{aligned} \qquad [3.2]$$

这里的 L 是对中的自由主义—保守主义分数。AC_{AA} 是一个虚拟变量（被告是美国黑人时赋值为 1，其他为 0），AC_L 也是虚拟变量（被告是拉美裔时赋值为 1，其他为 0）；JU_{AA} 是虚拟变量（陪审员是美国黑人时赋值为 1，其他为 0），JU_L 也是虚拟变量（陪审员是拉美裔时赋值为 1，其他为 0）。被告种族

的参照组是美国白人，而陪审员种族的参照组也是美国白人。尽管这个方程看上去很复杂，但用这种讨论原则其实很好操作。总体三向交互作用通过运用方程1.1的层级检验，即对方程3.2与去掉b_{14}、b_{15}、b_{16}和b_{17}系数相关变量项的模型进行比较。系数b_1是美国白人陪审员为美国白人被告定罪时，自由主义—保守主义倾向的作用。重新定义虚拟变量，重新计算方程，并分离出两个调节变量不同的参照组组合，就会产生9个斜率值相关统计，由3×3被告种族和陪审员种族表定义。在每一种情况下，关注点都是b_1。

任何包含L的双向交互项都反映了斜率之间的差异。例如，b_7代表的是陪审员为美国白人时，定罪分数对自由主义分数进行回归，被告为拉美裔的斜率均值减去被告为美国白人的斜率均值的差。一个给定的三向乘积项反映的是斜率之间差异的差。例如，b_{17}就是美国白人陪审员斜率差异减去拉美裔陪审员的斜率差异。尽管涉及很多系数，对它们的解释还是很简单的。

第3节 | 小结

研究者常对探索以三向交互作用为焦点的交互模型感兴趣。交互作用中的解释变量可能全是连续的,或是连续和定性变量的组合。在这两种情况下,研究者都宣称第一个变量为关键自变量,第二个变量是一阶调节变量,第三个变量是二阶调节变量。恰当的乘积项由此可被计算出来,然后交互模型包含了所有"主效应"项和所有乘积项。总体交互作用的显著性检验反映在与三向乘积项相连的回归系数里,并具有单一自由度。当总体交互作用超过一个自由度时,就需要运用层级检验比较三向交互模型和去除三向乘积项模型来评估总体的统计显著性。

非标准化交互效应的优势在于,回归系数反映的是单自由度的交互比较。标准化交互效应的优势在于其反映了三向乘积项的半偏相关系数平方。

对交互作用的理解通常关注与关键自变量相关的回归系数以及包含关键自变量各个乘积项的回归系数。针对不同的调节变量,我们使用不同的对中和再编码策略来帮助读者理解交互的本质。



第4章

其他重要问题

本章考察一系列与交互分析相关的事项，大部分内容都比前面讨论的更高深。

第 1 节 | 连续变量之交互作用的双线性本质

如前所述，当一个连续变量是交互作用的一部分时，重要的一点就是传统的包含乘积项检验的交互模型有一个特定形式，即双线性交互。

其他形式的交互也可操作，并且应当常规性地进行探索性分析，以保证交互作用的形式是正确的。在两个连续变量的情况下，经典乘积项方法反映了一个定义较窄，但适用性可能更宽的交互形式。如前所述，如果 X 是关键自变量，Z 是调节变量，那么乘积项方法会把 X 的系数当做 Z 的一个线性函数，但很可能 X 的系数变动是 Z 的非线性函数。如果那样，传统的乘积项分析就是设定有误的模型。探究这一问题的一个粗略但具有启发性的方法是使用分组差异回归(Hamilton, 1992)。在此情况下，调节变量(Z)按照差不多相等的数量被分成 5 到 10 个组，顺序分类。然后，给计算每一组中 Z 的均值或中位数，再对每一个 Z 组进行一次 Y 对 X 的回归分析。在由 Z 定义的 5 到 10 个组中，检查 Y 对 X 的回归系数，这揭示了根据每组 Z 的均值或中位数的变化，Y 对 X 回归系数是否存在增加或者减少的趋势。换言之，如果我们根据 Y 对 X 回归的系数与 Z 的每组均值(或中位数)作

图，就会存在一个线性趋势。如果并非这样，那么就需要一个更复杂的交互形式。

这种复杂的交互形式往往可使用乘积项的多项式回归。包含交互项多项式分析的导论，请参见 Jaccard、Turrisi 和 Wan(1990)。举个例子，X 和 Z 都是连续变量，要应用这样的模型，假设 X 的系数不是 Z 的线性函数而是 Z 的二次函数，步骤如下：(1) 确定关键自变量 X 和调节变量 Z；(2) 对 X 和 Z 进行必要的变换(比如对中操作)；(3) 计算调节变量的平方，Z^2；(4) 计算 X 和 Z 以及 X 和 Z^2 的乘积项；(5) 确定方程 $Y = \alpha + \beta_1 X + \beta_2 Z + \beta_3 Z^2 + \beta_4 XZ + \beta_5 XZ^2 + \varepsilon$。

进行层级检验可观察加入 XZ^2 项对模型的改进程度，表示这个二次项交互作用是否重要(或可简单地检查与 β_5 相连系数的显著性)。给定 Z 值，X 的系数是由 $\beta_1 + \beta_4 + \beta_5 Z^2$ 决定的。β_1 是当 $Z=0$ 时，X 的系数。可以变换 Z(见第二步)，使 Z 的 0 值代表一个理论上有意义的值，以此来分离在任意给定的 Z 值上，Y 对 X 的系数和置信区间。

对只涉及定性变量和连续解释变量的情况而言，假设 Z 是虚拟变量，取值为 1 和 0，表示属于哪个组。在此情况下，至少在一个组上(很可能在两个组上)，Y 都是 X 的一个非线性函数。拟合下面这个方程：

$$Y = \alpha + \beta_1 X + \beta_2 Z + \beta_3 X^2 + \beta_4 XZ + \beta_5 X^2 Z + \varepsilon$$

当 $Z = 0$ 时，X 对 Y 的作用反映在二次模型 $\alpha + \beta_1 X + \beta_3 X^2$ 中。为找出当 $Z=1$ 时 X 对 Y 的作用，必须重新给 Z 编码，使 Z 的参照组调转，重新计算乘积项，重新进行回归，并再次计算 $\alpha + \beta_1 X + \beta_3 X^2$。

有一些方法论学者（例如 Ganzach，1997）认为，假设简单线性关系和简单双线性交互关系的模型都过于有限，在研究中应当常规性地考察曲线关系效应。方法之一是拟合一个模型，允许 Y 和解释变量之间以及与交互项之间产生线性或二次的关系。同时，该模型会对 X 系数对 Z 的线性或者二次函数形式很敏感。对连续变量而言，这种模型会有如下形式：

$$Y = \alpha + \beta_1 X + \beta_2 X^2 + \beta_3 Z + \beta_4 Z^2 + \beta_5 XZ + \beta_6 XZ^2 + \beta_7 X^2 Z + \beta_8 X^2 Z^2 + \varepsilon \qquad [4.1]$$

在任意的给定 Z 值上，X 的作用都可以被分离出来，通过变换 Z 值策略，使 Z 的 0 值等于我们感兴趣的那个值，然后关注 β_1 和 β_2，分离当 $Z = 0$ 时，相应模型 $Y = \alpha + \beta_1 X + \beta_2 X^2$ 的系数。正如后面将讨论的，这种模型的优势是帮助理论家避免虚假相关或忽略真实的交互作用。批评者认为，这一方法可能过度拟合数据，且违背理论建构中的简洁原则。这两个观点都有道理。在方程4.1中，我们假设的是任何非线性的关系都是二次形式的。

第2节 | 计算关键自变量在调节变量取不同值时的系数

在前面的章节中,关键自变量系数都是在调节变量取不同值时计算出来的,即通过转换连续性调节变量或重新定义定性调节变量的参照组,然后在电脑上重新进行回归。

尽管这个方法有些繁琐,但其优势是可以产生所有我们感兴趣变量的标准误和置信区间。手工计算这些置信区间就不那么容易(相关方程见 Aiken&West, 1991)。当要计算原方程中的系数而不产生置信区间,不利用变换后的变量重新进行回归时,就会有问题。这一节将描述这么做的基本逻辑。让我们从方程 2.7 中的方法开始,再把这一逻辑推广到其他情境下。

请考虑 X 是关键自变量,Z 是调节变量的情况,方程是:

$$Y = \alpha + \beta_1 X + \beta_2 Z + \beta_3 XZ + \varepsilon \qquad [4.2]$$

要决定某些 Z 值上 X 的系数,首先必须分离出右边包含 X 的所有的项:

$$\beta_1 X + \beta_3 XZ$$

然后,我们把 X 提出来:

$$X(\beta_1 + \beta_3 Z)$$

产生的就是在任意给定 Z 值上的 X 值：

$$X\text{ 在 }Z\text{ 值上面的系数} = \beta_1 + \beta_3 Z \qquad [4.3]$$

在方程 4.2 中，如果 $\beta_1 = 1.2$，$\beta_3 = 0.05$，那么当 $Z = 2$ 时，X 的系数就是 $1.2 + (0.05)(2) = 1.3$。请注意，当 $Z = 0$ 时，方程 4.3 中的系数就是 β_1，这就表明 β_1 是以 $Z = 0$ 为条件的。

如果 X 和 Z 都是虚拟变量，那么方程 4.3 的逻辑依然成立，不过关注点就是相关虚拟变量了。例如，假设 X 有两个虚拟变量，而 Z 有两个虚拟变量，就产生了下面的方程：

$$Y = \alpha + \beta_1 D_{X1} + \beta_2 D_{X2} + \beta_3 D_{Z1} + \beta_4 D_{Z2} + \beta_5 D_{X1} D_{Z1}$$
$$+ \beta_6 D_{X1} D_{Z2} + \beta_7 D_{X2} D_{Z1} + \beta_8 D_{X2} D_{Z2} + \varepsilon$$

假设我们想分离出 D_{Z1} 和 D_{Z2} 都为 1 时，在 X 上 D_{X1} 为 1 的组与参照组的系数。首先分离直接包含 D_{X1} 的项和系数：

$$\beta_1 D_{X1} + \beta_5 D_{X1} D_{Z1} + \beta_6 D_{X1} D_{Z2}$$

然后把 D_{X1} 提出来，我们得到：

$$D_{X1}(\beta_1 + \beta_5 D_{Z1} + \beta_6 D_{Z2})$$

因此，

$$X\text{ 在 }D_{Z1}\text{ 和 }D_{Z2}\text{ 上的 }\beta = \beta_1 + \beta_5 D_{Z1} + \beta_6 D_{Z2}$$

在这个例子中，$\beta_1 = 0.2$，$\beta_5 = 0.3$，$\beta_6 = 0.4$，$D_{Z1} = 1$ 且 $D_{Z2} = 1$，那么 D_{X1} 的系数就是 $[0.2 + (0.3)(1) + (0.4)(1)] = 0.90$。[10]

三向交互作用使用相同的逻辑。对三个连续性解释变

量 X、Q 和 Z 而言，传统的交互方程是：

$$Y = \alpha + \beta_1 X + \beta_2 Q + \beta_3 Z + \beta_4 XQ + \beta_5 XZ + \beta_6 QZ + \beta_7 XQZ + \varepsilon$$

在给定 Q 值和 Z 值的组合时，X 的系数是：

X 在 Q 值和 Z 值组合上的 $\beta = \beta_1 + \beta_4 Q + \beta_5 Z + \beta_7 QZ$ [11]

而在给定 Z 值时，XQ 的系数是：

XQ 在 Z 值时 $\beta = \beta_4 + \beta_7 Z$

第 3 节 | 计算偏组成项

有时人们会说，乘积项在回归方程中代表交互效应，通过它们反映出乘积项是主效应和交互效应的混合体。只有当乘积项的组成部分与乘积项一起被包括在方程中时，此书描述的顺序关系才成立（与无限制的截距项一起）。排除一个或者更多乘积项的组成部分依然可能建立交互作用模型，但这就不同于我们这里所考虑的交互作用了。

传统的交互分析使用的是我们称之为“层级完整建构”模型。一个层级完整建构模型是指，所有的高阶交互项涉及的低阶组成项都被包含在模型中。例如，我们对 X 和 Z 之间的双向交互作用感兴趣，那么层级完整建构模型就应当包含解释变量 X、Z 和 XZ；如果我们对 Q、X 和 Z 之间的三向交互作用感兴趣，那么层级完整建构模型就应当包括解释变量 Q、X、Z、QX、QZ、XZ 和 QXZ；如果有一个定性解释变量（用虚拟变量 D_1 和 D_2 表示）和一个连续解释变量 Z，那么层级完整建构交互模型就应当包括 D_1、D_2、Z、$D_1 \times Z$ 和 $D_2 \times Z$。研究者使用层级完整建构模型时，就可用本书介绍的顺序关系了。

当然，也可以不使用层级完整建构模型，而对特定类型的交互作用建立模型。例如，一个简单的相乘模型可以采用

如下形式：

$$Y = \alpha + \beta_1 XZ + \varepsilon$$

这一模型(例如,乘积相关)的拟合受到 X 和 Z 度量的影响。层级完整建构模型使用的是变换,把 X 减去一个常数而不影响模型拟合,这在上面这个相乘模型中就不适用。如果一个度量在纯相乘模型中是随意的,那么这一模型的拟合也是如此。模型化非层级完整建构的交互作用需要熟练的技术才可运用。

第 4 节 | 变换

我们很依赖简单变换策略(测度值减去一个常数)来分离出简单效应以及相关的标准误、显著性检验和置信区间。这些变换可用于双向交互作用、三向交互作用、四向交互作用或更高阶交互作用。分析者简单地使用层级完整建构模型,并牢记系数的如下条件性本质:在任何时候,变量 X 只要涉及乘积项,那么与它相关的简单系数就都是以乘积项中其他变量取 0 值为条件的。与此类似,如果一个乘积项(XZ)涉及更高阶的乘积项(例如,XZQ),那么它与 XZ 相关的系数就都以更高阶乘积项中的其他变量取 0 值 ($Q = 0$) 为条件。

调节变量取某个值时,我们试图分离 X 对 Y 的简单效应,此时若使用变换策略,通常以原始分数来操作为最佳选择,即应用变换产生的 0 值对应我们感兴趣的那个值。有些研究者先对中 Z 分数,再应用变换策略来转换这些 Z 分数(减去或加上一个标准差)。尽管可以这么做,但根据经验,应用这种双重变换分数时,研究者往往会得到令人吃惊的结果。用代数来解释其背后的机制显得有点多余,我们在此提醒读者,最安全的方法是避免对解释变量进行双重变换,且总是从理论需要出发确定所需调节变量的值,通过应用变换策略在原始分数上减去一个常数,进而产生一个我们感兴趣的 0 值点。

第 5 节 | 多重交互作用

请考虑这种情况：研究者需要模型化结果变量 Y 作为三个连续变量 X、Q 和 Z 的函数。研究者不希望看到解释变量之间的三向交互作用，但试图评估所有可能的双向交互作用。在此，可使用多种策略。有些分析者使用组块检验，先拟合一个模型，把所有双向交互项都包含进来，再与一个没有任何交互项的模型进行比较(Kleinbaum, 1992)。利用方程 1.1 可进行这一检验。如果两个模型的拟合程度的差异微不足道，那么就说明没有任何一项交互作用是必要的，它们都可从方程中剔除。如果应用组块检验有明显差异，那么表明至少有一项交互项需要被保留。在此意义上，层级逆向消除策略可用于比较包含所有交互项模型与去掉某一特定交互项模型之间拟合程度的差异(通过方程1.1)。例如，研究者对评估 XZ 的交互作用感兴趣，那么就应当比较这一模型的拟合程度：

$$Y = \alpha + \beta_1 X + \beta_2 Z + \beta_3 Q + \beta_4 XZ + \beta_5 XQ + \beta_6 QZ + \varepsilon$$

与下面模型的拟合程度：

$$Y = \alpha + \beta_1 X + \beta_2 Z + \beta_3 Q + \beta_4 XQ + \beta_5 QZ + \varepsilon$$

如果这两个模型拟合程度差异很小，那就表明 XZ 项可

被剔除。然而，如果模型的拟合度差异是显著的，那么这一项就应被保留。

有些分析者用这种方式系统地评估每一个交互项。有些分析者就首先选择一项，如果这项被剔除，再评估余下的项（从方程中剔除已确定的可剔除项）。例如，首先检验 XZ，看它是否能够被剔除，并最终决定是否从模型中剔除它。当 XZ 已不出现在模型中时，再评估 QZ，关注点是逆向排除法。也就是说，我们评估的是：

$$Y = \alpha + \beta_1 X + \beta_2 Z + \beta_3 Q + \beta_4 XQ + \beta_5 QZ + \varepsilon$$

与如下模型相比：

$$Y = \alpha + \beta_1 X + \beta_2 Z + \beta_3 Q + \beta_4 XQ + \varepsilon$$

首先基于一些理论上的标准选择可能被排除的项进行评估。有时，我们首先评估全模型中相应系数具有最大 p 值的项，有时则两个标准兼顾。

在多元交互情境下，剔除某些项可以援引很多的模型拟合标准，对相关策略的争论持续存在。深入考察相关情况已经超出本书的范围。感兴趣的读者可参考 Bishop、Feinberg 和 Holland(1975)、Hosmer 和 Lemeshow(1989)以及 Jaccard (1998)的著作。需要提醒读者的是，这些看似“怪异”的情况，在顺序考虑多元交互项时常会发生。例如，组块检验可能提示说至少有一个乘积项应被保留，但对具体每一项的评估却说每一项都应被剔除出模型。或者，单个项的检验证明某一项应被保留，其余所有项都应被剔除，但当其余项被剔除后，原先的这个保留项也变得不显著，只剩下微弱的解释力。那么，该如何处理这种情况呢？这取决于要回答的理论

问题、研究者的统计框架(例如,虚无假设检验、数量估计、置信区间估计)以及数据的形式。在大部分情况下,选择剔除项是直接的、无争议的,但并非所有情况都如此。

当回归方程只包括两个单独的交互项时(比如,三个连续变量 Q、X、Z 和乘积项 XZ、QZ 都保留在方程中,不包括其他交互项),其理解与前文给出的交互项的系数理解一样,只不过前文中的系数是其他双向交互作用(以及其他的协变量)保持不变的情况下得到的。因此,低阶项系数以所有乘积项涉及的其他变量取 0 值为条件。

第6节 | 标准化与非标准化系数

本书讨论的回归系数通常指与解释变量相关的非标准化回归系数。尽管交互分析中可使用标准化系数，但这些系数可能会使研究者误入歧途，且不符合我们提出的规则。我们一般不建议使用标准化系数，尽管在有些情况下使用它们很正当。举一个标准化回归系数局限性的例子，假设一个简单的双变量回归，即收入对受教育年限进行回归，我们希望能够发现受教育一年的“价值”。这一分析在两个不同的种族群体间展开，即美国黑人和美国白人。假定分析发现，两个群体具有相同的标准化回归系数，表明教育变动一个标准差，收入会变动 0.50 个标准差。研究者或许就此得出结论说，在两个群体中，教育的“价值”是相同的。假设两个群体教育的标准差都是 3.0，但美国白人的收入标准差是 15000，而美国黑人的收入标准差是 6000，那么问题在于，美国白人的非标准化系数是 2500，而美国黑人的非标准化系数是 1000 美元，即对美国白人而言，一年的额外教育预计将增加 2500 美元的收入，而对美国黑人而言，一年的额外教育只增加 1000 美元的收入。两个组之间存在明显的差异，而标准化回归系数不能反映这一点。[12]

标准化分析的问题在于，它对两个群体产生不同的度量

单位。因为标准差大小不同，所以标准差这一度量就不同。对美国白人而言，这个度量是15000美元；对美国黑人而言，这个度量是 6000 美元。在不同的度量上比较这两个群体，就像是以美元为单位度量一个群体的收入，而用英镑度量另一群体的收入。关于标准化系数的其他限制，请参见 Jaccard、Turrisi 和 Wan(1990)的著作。

第7节 度量性质

度量性质在交互分析中很重要。通过经典的潜变量模型来理解测度问题是最好的。建构的可观测测度可被看做潜变量的一个指标,代表的是有关变量的真实建构。在实践中,我们不知道某人的真实分数,但可利用可观测测度去估计它。可观测测度不仅受某人真实的潜变量值的影响,还受测量误差的影响。图4.1表示这一模型的路径图。如果测量的是可观测测度和潜变量之间的关系,那么图4.1代表的是一个回归方程,其中,可观测测度对潜变量进行回归:

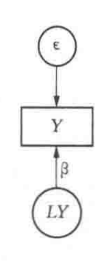


图4.1 一个测量模型的路径图

$$Y=\alpha+\beta LY+\varepsilon$$

上面这个测量模型描述了一个可观测测度如何映射在潜变量上。在这个测量模型上,不同的组在截距、斜率或误差的方差上可能都不同。这些差异会影响在真实的潜变量上群组之间差异的推论(但这却是我们的研究兴趣所在)。例如,两个组在潜变量上具有相同的均值分数,但它们在测量模型中的截距不同,观察变量的平均分数也不同。如果两个组在测量模型中的回归系数也不同,那就说明即使它们在

潜变量上有相同的方差,但在观察到的测度上也可能具有不同的方差。如果两个组误差的方差不同,那么测度就依赖于这些组(其他条件保持一致)。最理想的情况是,不同的组由不同的调节变量定义,具有相同的截距和斜率,在测量模型中,误差的方差为 0 或者接近于 0。偏离这一理想状态会导致交互分析的问题。有关度量影响的讨论、检验度量一致性的方法、调整度量的方法,参见 Vandenberg 和 Lance(2000)以及 Busemeyer 和 Jones(1983)的著作。

有一些研究者假定这里讨论的交互分析只对比例尺度适用,但事实并非如此。这种方法可以被有效地运用于定距测度或是与定距测度的特点相当接近的测度。后文会解释这一点。

有一些研究者错误地把“尺度”当做定距或定序的。我们必须认识到,度量的性质与尺度不一致而与数据一致,因此它受数据收集过程中所有方面的影响。一组测度具有定距属性的程度不仅依赖于用于观察的尺度,还受产生观测样本的特定个体、收集时间、数据收集情境等因素的影响。我们来考察一个简单的教学案例,即用两个不同的度量——英寸和高度顺序——来测量五个人的身高。

个　体	高度(英寸)	身高顺序
A	72″	5
B	71″	4
C	70″	3
D	69″	2
E	67″	1

众所周知,利用英寸的测度具有定距属性。例如,两个

分数之差为1,对应在身高这个维度上有同样的物理差异。A和B的真实身高差异与C和D的真实身高差异是相同的,度量的性质反映了这一点(例如,72－71＝1,70－69＝1)。同样,D和E的差异是69－67＝2,A和C之差异也是2。这些差异也反映了身高维度上相同的真实差异。请注意,这些属性并不适用于定序测度。A和B的差异是1(5－4＝1),而D和E的差异也是1(2－1＝1)。这个相同的(顺序)差异对应了不同程度的真实身高维度的差异(用英寸进行测度时,D和E之间真实的身高差异明显比A和B之间真实的身高差异大)。对这些个体而言,定序测度只有定序属性,没有定距属性。

我们再用下面这些分数考察五个不同的个体:

个　体	高度(英寸)	身高顺序
A	72″	5
B	71″	4
C	70″	3
D	69″	2
E	68″	1

请注意这五个人,定序测度同样具有定距的属性。个体A和B的差异为1,这与个体D和E的差异相同。这些差异所对应的真实物理维度上的差异具有相同的量。在此例中,传统上认为的定序"尺度"事实上产生了一个具有定距属性的测度。假设E的身高不是68英寸,而是67.9英寸。那在此情况下,定序测度在严格意义上就不是定距的。如果它们是没有相反效应的定距层次,那么它们就接近于或可被当做定距来处理。

这个例子揭示了,最关键的问题不在于一组测度是定距还是定序,相反,重要的是一组测度在何种程度上接近定距特征。如果这种相似性非常高,那么就可以假设它们具有定距属性,并利用统计方法对数据进行有效分析;如果这种相似性很低,那么就需要另外的分析策略。在此情况下,交互分析可通过定序层次的数据以及它们所接近的定距层级特征来进行。

第8节 | 测量误差

在有关交互作用的统计文献中，大家相当关注一个话题，那就是测量误差带来的偏差。众所周知，在多元回归中，不可靠的测量会产生有偏的回归系数(Bohrnstedt & Carter, 1971)。社会科学家经常使用有误差的测度进行研究。因此，测量误差是交互作用分析的潜在问题，尤其在涉及连续变量时。

Busemeyer 和 Jones(1983)使用经典的检验理论表明，测量误差具有减弱乘积项层级估值的效果。减弱程度是乘积项可靠性的函数，我们称之为 Pr。在标准的统计假定下，减弱的量反映在主效应模型 R^2 的变化上，为 $Pr(R_2^2 - R_1^2)$。例如，如果增加一个乘积项，真实的解释方差增量 $(R_2^2 - R_1^2)$ 为 0.20，而乘积项的可靠性为 0.70，那么，观察到的解释方差将会是 $(0.20)(0.70) = 0.14$，其他条件则保持不变。

在合理的统计条件下[13]，如果 X 和 Z 之间的真实相关性为 0，那么乘积项 XZ 的可靠性就等于 X 的可靠性乘以 Z 的可靠性。因此，若一个测度是相对可靠的 $(r_{XX} = 0.80)$，第二个测度是相对不可靠的 $(r_{ZZ} = 0.50)$，那么乘积项的可靠性就会比可靠性最低的组成项的可靠性还低，即 $(0.80)(0.50) = 0.40$。当 X 和 Z 的真实相关性增大时，乘积

项的可靠性也会增大。但根据社会科学中通常观察到的相关性和可靠性可知，这并不会增加太多乘积项的可靠性(Busemeyer & Jones, 1985)。这些事实强调了测量误差给检验交互作用带来的困难，特别是在低统计解释力的情况下。对假设检验而言，利用大样本可以抵消测量误差所降低的解释力，但一个大样本不必然消除测量误差导致的回归系数的偏差(例如，Busemeyer & Jones, 1983; Evans, 1985)。

有几个处理测量误差问题的解决方案。Cohen 等人(1975)、Bohrnstedt 和 Marwell(1978)、Heise(1986)、Fuller 和 Hidiroglu(1978)建议的方法都要求预先对组成变量的可靠性有所了解。Cohen 等人(1983: 410)发现，他们在 1977 年的书中提出的修正方法是错误的，那个方法倾向于高估回归系数。Bohrnstedt和 Marwell 的方法有几个局限，具体见 Busemeyer 和 Jones(1983)的著作。Heise(1986)发现，他的方法只在一开始可靠性就相对较高的情况下才能得到令人满意的效果(比如，高于 0.90)。Fuller 和 Hidiroglu 的方法很有前景，但只针对那些不含乘积项的模型。若把这种方法扩展到乘积项进行分析，将会很有价值。

第二组处理测量误差的方法是使用潜类变量结构方程模型(有关这类方法的导论，见 Jaccard & Wan, 1996)。这些策略依赖于每个真实变量建构都有多个指标，把误差理论整合进模型检验和参数估计中。使用此类方法的交互模型受到极大的关注。这类方法的问题是需要大样本，并要假定解释变量都是正态分布的。有几个策略可以放松正态分布假定。这包括由 Bollen(1996)、Bollen 和 Paxton(1998)发展出来的二阶段最小方差方法(2SLS)、准最大似然估计法(Klein &

Muthen，未刊稿）以及含误差因子分数法（Wall & Amemiya，2000）。目前并不能判断哪个分析方案必然优于另一个分析方案。尽管还有许多工作需要做，但这类方法依然是有前景的。

依靠传统多元回归方法的分析者在分析交互作用时，必须承认存在由测量误差引起的潜在偏差。分析者应当使用有效和可靠的度量，并在作结论时保持必要的警惕。忽视测量误差等于假定有完美的可靠性，这意味着社会科学在进行复杂的理论检验前，应花费相当多的时间和精力来发展高质量的测度。心理计量学、问卷设计和精神物理学中存在大量关于降低测量误差建议的文献。对这些实践的有用讨论，请参见 Anderson（1981）和 Wegenar（1982）的著作。

第9节 | 稳健分析与假定的违反

到此为止，我们的分析都假定有关OLS回归推论的标准成立。我们也预设不存在掩盖数据根本趋势的极端值。

很不幸，这些OLS回归的假定常常被违反，推论检验会相应地被弱化。应对假定被违反的一个方法是使用初步检验来评估这些假定的有效性(比如，正态分布检验或方差异质性检验)，当发现假定被违反时，应修改相应的分析策略。许多此类检验都缺乏统计解释力，除非使用大样本，否则通常很难发现非正态分布或方差异质性。另外，传统OLS的t检验和F检验并非是在利用“筛选检验”来决定是否继续分析的情况下发展起来的。应用这种筛选，可能改变样本的内在分布，弊大于利。

另外一个策略就是以某种方式变换数据，使统计模型拟合数据。有时候，这个变换解决了一个问题(非正态)，同时产生了其他问题(方差异质性)。有些变换改变了变量单位，使新度量不具备真实世界的意义。从实践角度来看，这使测度问题变得难以处理。最后，通常的变换哲学是理顺本末关系。与其通过变换来操作数据以符合潜在的统计模型，何不直接使用一个适合于已有数据的统计模型呢?

20多年前，在很多情况下，发现这样一个统计模型或许

会被证明是不可能的。但随着电脑技术的迅速发展,稳健统计方法领域有了长足的进步。现在有切实可行的分析方法,并且它们不像传统统计方法那样要求强假定,这就产生了非常好的比较统计力量,且不用回避极端值。有些方法使用了本书讨论的原则(比如,乘积项产生同类型的系数理解),却是在不回避极端值且产生稳健标准误和稳健置信区间的情况下。有些方法涉及完全不同的框架,比如基于平滑化的方法。关于这些稳健方法,请参考 Wilcox(1997、2001)的著作。Wilcox(1997)特别介绍了关于某些形式交互作用的稳健分析。

最新进展是把稳健方法整合到结构方程模型中,同时涵盖违反假定、极端值和测量误差。这些方法包括标准误、置信区间引导估计以及共同关注不需假定分布的估计(Arbuckle & Wothke, 1999)。这些技术拥有广阔的前景。

第10节 | 被试者内和重复测量设计

有时，研究者要求在被试者内或是重复测量设计情况下，检验斜率之间的差异。设计之一是在两个时间点上测量结果变量，然后对两个时点共同的、稳定的解释变量进行回归。例如，分别用受测试者在七年级时和九年级时与父母关系的满意度对性别进行回归。我们感兴趣的是性别对关系满意度的作用在七年级时是否与在九年级时不同。这种设计产生了下面两个方程：

$$Y_{t1} = \alpha_{t1} + \beta_{t1} X + \varepsilon_{t1}$$
$$Y_{t2} = \alpha_{t2} + \beta_{t2} X + \varepsilon_{t2}$$

这里，X 是随时间不变的解释变量，Y_{t1} 是在时点 1 测量的结果变量，Y_{t2} 是在时点 2 测量的结果变量。Judd、Kenny 和 McClelland(2002)表明，如果假定图 4.2 的路径模型成立，那么两个方程的斜率相同的虚无假设检验就可通过传统的 OLS 回归来进行。假设 Y_{t1} 和 Y_{t2} 对 X 的差异预测如下：

$$Y_{t1} - Y_{t2} = a + bX + e \qquad [4.4]$$

方程 4.4 中的斜率 b 等于 b_{t1} 和 b_{t2} 的差异，检验 b 的显著性可以评估虚无假设 $\beta_{t1} = \beta_{t2}$。Judd 等(2002)讨论了将这种检验扩展到多于两个重复测度的情况。

在第二类被试者内设计中，设计关注 X 和 Y 都随时间变动的个案，产生的两个方程是：

$$Y_{t1} = \alpha_{t1} + \beta_{t1} X_{t1} + \varepsilon_{t1}$$
$$Y_{t2} = \alpha_{t2} + \beta_{t2} X_{t2} + \varepsilon_{t2}$$

虚无假设 $\beta_{t1} = \beta_{t2}$ 即被称为“序列调节”的检验。James 和 Tetrick(1984)描述了一个 T^2 统计量，基于最小方差回归模型来检验序列调节，假定图 4.3 中的模型成立。

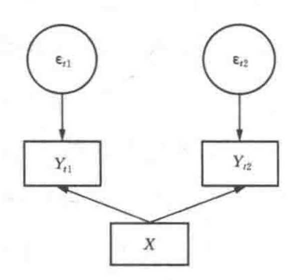


图 4.2 具有稳定解释变量的纵贯模型

这些程序的一个局限性是图 4.2 和图 4.3 中的内在模型可能不正确，这会导致检验失效。检验中经常存在相关误差，或者说，在时点 1 上的结果变量可能对时点 2 上的结果变量在 X 之外还有独立的作用。对图 4.3 中的模型而言，Y_{t2} 或许不仅受 X_{t2} 的影响，还受 X_{t1} 独立于 X_{t2} 的影响。如果这个模型是设定错误的，那么这些参数差异的检验就被低估了。检验交互作用假设最好的方法是结构方程模型(SEM)(Kline, 1998)。SEM 可以容纳较多的因果模型和误差结构，在此检验情境下，它们经常要求中等样本或大样本，因为它们依赖于渐进理论。计量经济学中的特定方法对此有所帮助(Greene, 1997)。

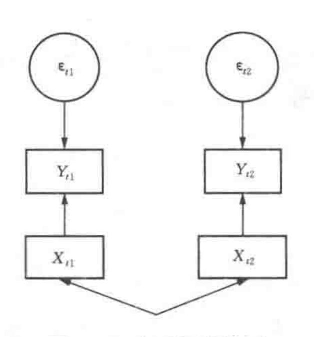


图 4.3 序列调节模型

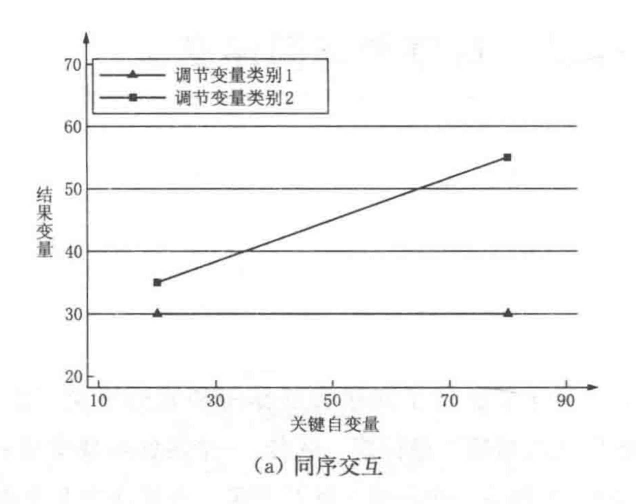
另一个常用的、基于时间的交互分析是增长曲线模型。在此方法中，研究者指定一个数学函数来描述个体的一个变量如何随时间的变化而变化。例如，某人的阅读能力(Y)以1、2、3、4、5 和 6 表示，可能随年级的变动而变动，且这种变动是线性的。可用斜率和截距对这个个体的线性变动进行描述，斜率表示的是个体时间变化一个单位，阅读水平会变化多少个单位。如果时间单位是年，那么这些斜率就表示从这一年到下一年，阅读水平有多少变化(从 1 到 6)。这些斜率可能会有个体差异，有些孩子展示出更陡的斜率，表明从这一年到下一年，阅读水平有更大的变化；另外一些学生则会展示出相对平缓的斜率，表明从这一年到下一年，阅读能力的变化比较小。我们感兴趣的是预测斜率的量级。例如，平均而言，私立学校的孩子相比公立学校的孩子，是否会有更陡的斜率？那个斜率的量级是不是孩子社会经济出身的函数？这些问题(都基于交互作用)可用在层级线性模型统计文献中发展出来的增长曲线模型来解释(Bryk & Raudenbush, 2002)。增长曲线模型可用结构方程模型框架加以参数化(Duncan, Duncan, Strycker, Li &Alpert, 1999)。

第11节 | 同序和非同序交互

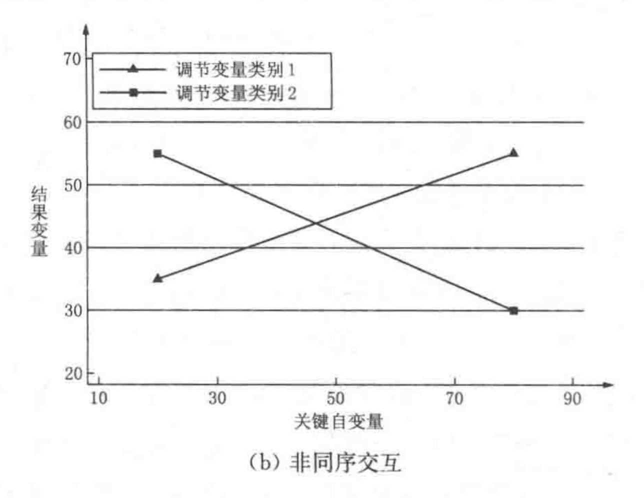
社会科学家区分了同序交互作用和非同序交互作用。我们通常在此情境下进行这一区分：一个定性解释变量有二到三个组，另外有一个连续性解释变量。非同序交互是指，Y在一个组中对连续变量进行回归的回归线与Y在另一个组中对连续变量进行回归的回归线相交，这也被称为“交叉交互”。同序交互是指两组回归线是非平行的，但并不相交。图4.4展示的就是非同序和同序交互作用。

统计学家已经表达了关于交互作用的谨慎态度。他们认为，这种交互可能是因变量度量尺度的加工品。通过Y分数的一个单调变换，非平行的回归线常常可以被变成平行的。如果Y的度量间距是随意的，那么从科学简洁性的角度来看，执行这种变换以去除虚假的调节作用就是合理的。尽管如此，如果其度量是有意义的，那么同序交互作用就不应被撤销。正如Cronbach和Snow(1981)指出的，这种交互作用从本质上来说很重要，当与成本—收益分析结合时，可对分类决策起关键作用。

对任何一组非平行线而言，总存在两线相交的点。在此意义上，理论上来讲，所有交互作用都是非同序的。如果在研究的分数取值范围内(例如，IQ分数在90和110之间)，回



(a) 同序交互



(b) 非同序交互

图 4.4　同序和非同序交互作用

归线不相交,交互作用即被认为是同序的。请考虑两个组的例子,每个组都可被描述为 Y 对 X 回归的线性方程。使用下列方程,可识别回归线对两个组的相交点是连续变量的哪个值:

$$P_1=(a_1-a_2)/(b_2-b_1)$$

这里，a_1 是第一组结果变量对连续解释变量回归的截距，a_2 是第二组结果变量对连续解释变量回归的截距，b_1 是第一组结果变量对连续性解释变量回归的系数，b_2 是第二组结果变量对连续性解释变量回归的系数。

为了展示相关计算和实际应用，请考虑下面这个例子：一位心理学家对两种不同类型的、针对儿童自尊的疗法的相对作用感兴趣。研究者设计了一个实验，旨在通过干预增强孩子的自尊心。一半孩子接受干预 A，另一半孩子接受干预 B。分组是随机的。研究没有控制组，因为先前的研究表明，与无干预情况相比，这两种干预方法都能有效地增强孩子的自尊心。孩子对这两种干预的反应被认为受到其与父母关系质量的调节作用。对与父母具有相对积极关系的孩子而言，干预方案 A 比干预方案 B 更有效；对与父母关系相对较差的孩子，干预方案 B 比干预方案 A 更有效。接受干预的类型采用虚拟编码，使用单一虚拟变量；孩子与父母关系的质量基于一位临床医师的评分，采用一个 51 分的尺度（从 0 到 50）。这位临床医师与孩子及其父母相处了相当长的时间，并在访谈基础上给出评分，分数越高，关系质量越好。再利用一个 0 到 20 的尺度测量自尊心，分数越高，自尊水平越高。在此研究中，干预后的自尊心是结果变量，干预类型是关键自变量，而父母—孩子关系则是调节变量。自尊心分数对虚拟变量、关系质量的测度以及这两个变量的乘积项进行回归（此例未进行任何均值对中操作）。交互作用是统计显著的。

首先，我们计算两个组单独的回归方程，自尊心对关系质量分数进行回归。这可运用第 2 章中讨论的程序进行

分离：

$$干预A: Y = 1.193 + 0.098Z + e$$
$$干预B: Y = 7.193 + (-0.107Z) + e$$

用一张回归线的图就可表明交互作用是非同序的。两线的相交点是：

$$P_1 = (1.193 - 7.193)/(-0.107 - 0.098) = 29.27$$

此点是 Z 的分数，两组的自尊心分数相同。因此，当关系质量分数等于29.27时，干预后自尊心分数对两个干预条件来说是相同的；当关系质量分数高于 29.27 时，在干预 A 下的干预后的自尊心分数就会比干预 B 下的更高；当关系质量分数低于 29.27 时，干预 B 下的干预后的自尊心分数就会比干预 A 下的更高。

这个分析结果提供了孩子该接受哪种干预方案的建议。如果在干预前，孩子与父母的关系分数低于 29.27，那么他们应当被施以干预 B；如果在干预前，孩子与父母的关系分数高于 29.27，那他们应当被施以干预 A。

Cronbach 和 Gleser(1957)回顾了这些处理的逻辑以及教育、组织和心理学研究中的分类决策。作者提出，受测试者分配到处理组的过程(例如，临床干预、教育课程、工作类型)经常受到非同序交互确定的相交点的指引。那些在相交点右边的人被分到一个处理组中，而交叉点左边的人被分到另一处理组中。相反，同序交互建议对所有人使用同样的处理。

第12节 | 显著性的区间

尽管上面的相交点分析很有用，但我们必须承认，在确定相交点过程中存在着样本误差。Potthoff（1964）基于Johnson和Neyman（1936）的经典之作，发展了一套方法，他在非同序交互中建立了一个与交叉点理解相关的"显著性区间"。这一方法关注研究者有一个定性关键自变量、一个连续性调节变量。对两组个案而言，这一方法定义了调节变量的一个取值范围、一个组的成员相比关键自变量其他组的成员，不会有更高的 Y 分数。该方法已经考虑了样本误差（Aiken & West，1991）。

在前文的例子中，应用下面的描述方法产生了一个值域，即27.36到31.07。这表示，当父母—孩子关系质量超出31.07时，可以很有信心地说（基于95%的置信度），接受干预A的儿童，其干预后的自尊心会超出那些接受了干预B的儿童。当父母—孩子的关系质量低于27.36时，同样可以很有信心地预测，接受干预B的儿童的干预后的自尊心会超过接受干预A的儿童。

相关的值可以通过计算下面这两个值（我们称之为 CV）来获得：

$$CV = [-B \pm (B^2 - AC)^{1/2}]/A \qquad [4.5]$$

为了定义 A、B 和 C 的值，使 Y = 结果变量，Z = 连续性调节变量，N = 包括所有组的关键自变量上有取值的个案总数，n_1 = 组 1 的个体数量，n_2 = 组 2 的个体数量，F_α = 具有两个和 $N-4$ 个自由度的 F 值表，对应于实验者定义的关键 F 值，并与一个预先设定的 α 水平相连（一般是 0.05），M_1 = 组 1 的 Z 均值，M_2 = 组 2 的 Z 均值，S_1 = 组 1 对 Z 回归的平方和，S_2 = 组 2 对 Z 回归的平方和，E = 在组 1 上 Y 对 Z 回归的残差平方和加上在组 2 上 Y 对 Z 回归的残差平方和，a_1 = 组 1 上 Y 对 Z 线性回归的截距，a_2 = 组 2 上 Y 对 Z 线性回归的截距，b_1 = 组 1 上 Y 对 Z 线性回归的斜率，b_2 = 组 2 上 Y 对 Z 线性回归的斜率。那么，

$$A = [-2F_\alpha/(N-4)]E[1/S_1 + 1/S_2] + (b_1 - b_2)^2$$

$$B = [2F_\alpha/(N-4)]E[M_1/S_1 + M_2/S_2] + (a_1 - a_2)(b_1 - b_2)$$

$$C = [-2F_\alpha/(N-4)]E[N/(n_1 n_2) + M_1^2/S_1 + M_2^2/S_2] + (a_1 - a_2)^2$$

组内回归方程不应当使用 Z 对中的度量。对这种方法、相关电脑程序以及对更复杂情况的扩展和更加深入的讨论，请参见 Aiken 和 West（1991）的著作。Cronbach 和 Snow（1981）讨论了这些方法的优点和局限性。

第13节 | 混合交互

有些分析者提出，交互作用经常与 X 对 Y 的曲线作用混合在一起，因此使交互分析变得复杂（Lubinski & Humphreys, 1990）。数据可能是 X 和 Y 之间曲线关系的一个结果，但当一个交互模型使用 X、Y 和另一个变量 Z 来拟合数据时，很可能会得到一个虚假的交互作用。

为了展示基本想法，在此举一个例子，Y 和 X 的内在关系模型是曲线的二次项关系，可用下面的模型表示：

$$Y = \alpha + \beta_1 X + \beta_2 X^2 + \varepsilon$$

而交互模型则表示为：

$$Y = \alpha + \beta_1 X + \beta_2 Z + \beta_3 XZ + \varepsilon$$

结果是，当 X 和 Z 的相关性增强时，X^2 和 XZ 的相关性也增强了。这意味着，若 X 和 Z 相关，将会在两个模型之间产生混合（Busemeyer & Jones, 1983; Lubinski & Humphreys, 1990）。针对二次项的本质，内在的关系是曲线的，但拟合一个（线性）交互模型会导致关于交互作用的虚假结论。如果交互和曲线的作用都是可操作的，那么拟合传统交互模型会导致错位交互、虚假的交互或是误导的交互作用，与真实交互作用的正负号都是相反的（Ganzach, 1997）。

问题的实质在于拟合了设定错误的模型到数据上，这么做使理论家误入歧途。研究者需要仔细考虑可能适合数据的模型，然后再探索这些相应的模型。如果一个曲线作用在理论上是不可能的或不具备概念合理性的，那就不应被使用。如果一个模型理论上有效，那就应被考虑。可能的情况是，当协变量取曲线作用并代入方程时，交互项就变得统计不显著。这不一定代表交互模型失效了，它只表明存在一个备择模型可解释这些数据。我们的底线是运用常识和理论来组织模型的类型，研究者必须认识到需要考虑多重模型，并在确定交互模型前进行比较。Lubinski 和 Humphreys (1990)及其他人的著作建议研究者在关注交互作用时，考虑曲线模型作为竞争性备择模型的有效性。

第 14 节 | 最佳实验设计和统计解释力

有些分析者哀叹，由于统计解释力低，在线性模型中发现交互作用非常困难。McClelland 和 Judd(1993)探索了大量在实际情境中引发低统计解释力的原因。这些研究者指出，利用统计解释力来发现交互作用，很大程度上依赖于乘积项组成变量的分布本质(例如，X 和 Z 是乘积项 XZ 的组成部分)。McClelland 和 Judd(1993)建议，当实际限制导致小样本和交互作用的低统计解释力时，可对 X 和 Z 的极端个案进行过度取样。然而必须小心地使用这些方法，因为它们会产生有偏的标准化作用大小的估计，例如在比较加叠模型(只含主效应的模型)和交互模型(全模型)复相关系数平方的增量时。

对交互作用进行统计解释力分析时，要注意交互作用大小的理论限制。例如，对交互项传统的解释力分析基于在“只含主效应”模型中加入一个乘积项，并观察可额外解释多少方差。如果理论表明，在两个解释变量中存在先验的同序交互，且在总体中这种交互作用不为 0，那么根据定义，含主效应的总体模型的复相关系数平方必然不为 0。同序交互作用越大，主效应模型中的效果也越大，这会被当做基准模型来评估交互项的统计解释力。Rogers(2002)描述了这两个

作用大小之间的正式关系。对同序交互作用进行解释力分析时,必须尊重这种关联,因为应用解释力分析时,研究者往往会不经意地选定一个理论上不可能的总体标准化交互作用大小。更多细节请参见 Rogers(2002)的著作。

第15节 | 协变量

协变量可被加入到本书讨论的任一回归模型中，这并不改变对相关交互作用的理解。例如，一个协变量 Q 可以被加入到乘积项模型中：

$$Y = \alpha + \beta_1 X + \beta_2 Z + \beta_3 XZ + \beta_4 Q + \varepsilon$$

在此模型中，β_3 是保持 Q 不变，Z 变化一个单位导致 X 对 Y 作用变动的单位。截距是当 X、Z 和 Q 都等于 0 时，预测的 Y 的均值。对 Q 进行简单线性转换不会影响其他解释变量的系数，但当 Q 的 0 值点改变后，就会影响截距项。

第 16 节 | 控制实验误差

在有些交互分析中，研究者进行的是多个单自由度比较。出现这种情况时，对一个给定的比较而言，作为标杆的 alpha 保持在一个特定值上(通常为0.05)，但至少在一组比较中，发生类型 1 错误的概率大于标杆 alpha。在此情况下，有些研究者就援引统计调整来维持一个特定的比较水平——实验误差的alpha水平(例如，在一组比较中至少有一个发生类型 1 错误的概率)。最流行的方法是传统的 Bonferroni 程序，尽管这一技术是保守的。更有力的备择方法是有效地运用类型 1 错误比例(Westfall et al., 1999)。例如，Holm(1979；亦见 Holland & Copenhaver, 1988；Seaman, Levin & Serlin, 1991)建议用一系列的修订版的 Bonferroni 模型。下面将介绍它们的应用。首先，通过每一组比较，在比较的组中得到一个 p 值。将 p 值从大到小排列。如果 p 值相同，那么它们就随意排序或利用理论标准排列。最小的 p 值与 $0.05/k$ 的 alpha 值进行比较，k 是这个组比较的数量。如果这个检验导致拒绝相关虚无假设(因为观察到的 p 值比调整的 α 值小)，那么倒数第二小的 p 值就与另一个 alpha 水平 $0.05/(k-1)$ 进行比较，$k-1$ 是剩下的比较数量。如果这个检验导致拒绝虚无假设，那么再次小的 p 值就与

$0.05/(k-2)$ 进行比较，如此进行，直到发现差异不显著为止。一旦观察到差异统计不显著，那么剩下的比较都可被视为不显著的。

如果报告的单自由度主要是展示性的，用于向读者提供一个交互作用的含义，那么就不必援引实验控制。如果研究者将基于分析作出理论陈述，此问题就更为重要。援引实验误差控制是复杂的，且会受一系列条件的影响，包括统计解释力和类型 1、类型 2 错误。

第 17 节 | 整体检验和交互作用

交互分析的常见策略是先对交互作用执行整体检验，只有当整体作用统计显著时，才可进行单自由度的交互比较。整体检验作为单自由度交互比较的基础，可以保证研究不会扩大类型 1 错误。一般而言，大部分二步法都不足以作为控制实验误差比例的方法(Jaccard，1998；Wilkinson，1999)。另一个策略就是直接从具有理论意义的单自由度比较入手，独立于整体检验结果来援引控制实验误差的方法(修改的 Bonferroni 检验)。这并不意味着交互作用的整体检验没有意义。若要反映两个或更多变量的总体交互作用的大小，这种检验就很有用。另外，如果整体交互作用离统计显著性很远，此时就不可能有任何的交互比较会显著。因此，整体检验是省力的工具。关于此问题的进一步讨论，参见 Jaccard (1998)的著作。

第18节 | 常见的误用

在研究文献中，常有一些差劲的研究与交互分析相关。我们前面已经提到了一个，它为两个或更多组单独计算回归方程，未对这些差异进行正式检验就宣称组间存在差异。另一个例子是交互分析以中位数为临界点，把一个连续解释变量降格为一个两值指标。这一策略常常被用来使传统方差分析法可进行交互作用分析。这种实践是不可取的，因为它们忽略了有用的信息，常常导致更低的统计解释力，并且它们引入了虚假作用(Maxwell & Delaney, 1993)。运用普通的线性模型可以有效地分析连续变量和定性变量之间的交互作用，而无需求助于降格连续解释变量为粗略的、只有两个值的指标。

第19节｜聚类数据和随机系数模型的交互作用

有些研究设计涉及聚类数据，研究者感兴趣的是聚类效应对 Y 在 X 上的斜率的影响。例如，研究者收集了5000个学生的数据（有50所学校，每所抽一个100学生）。研究者感兴趣的是同侪压力（X）对吸毒（Y）的影响以及这种作用是否与学生所在学校的规模（Z）有关。执行此分析的方法是，首先记录给定个体的 Z 值（学生所在学校的规模，因此来自同一所学校的学生有相同的 Z 值），然后形成乘积项 XZ，再进行OLS分析，用 Y 对 X、Z 和 XZ 进行回归，$N=5000$。这一策略是有问题的。限制之一是总体的残差分布可能不是独立的。来自同一所学校的学生相比其他学校的学生，彼此更为相似，这种机制会导致引入残差的相依性。[14] 因此，进行聚类分析必须考虑这一问题，而上面的回归策略并没这么做。

在此情境下，统计学家经常用一个不同于传统OLS的统计模型，即随机系数回归或者层级线性回归。比如，我们认为回归系数在不同学校之间的差异可用下面的方程来模型化：

$$\beta_j = \alpha + \beta Z_j + \varepsilon_j \quad [4.6]$$

这里的 β_j 指的是在学校 j 的 Y 对 X 的回归系数，Z_j 是学校 j 的规模大小，α 是 β_j 对 Z_j 回归的截距项，β 是 β_j 对 Z_j 回归的斜率，ε_j 是 β_j 对 Z_j 回归的残差项。总体而言，ε_j 被假定为正态分布，均值为 0，在给定 Z 值上有固定的方差。在方程 4.6 中出现 ε_j 是随机系数回归模型的定义规定的。这样一个模型与传统的伴随乘积项的回归模型不同，方程 4.6 要求用特定的分析模型来估计 α 和 β 并检验估计值的显著性。随机系数回归模型本身关注交互作用（因为 X 对 Y 的作用随 Z 的变化而变化），但它们通常应用于聚类数据。对这些方法的讨论，参见 Bryk 和 Raudenbush(2002)的著作。

第20节 | 连续还是离散解释变量

本书中有些例子使用了多值的、定量的、离散的变量作为解释变量，并且把它们当做连续变量处理。当一个解释变量有多个值并基本符合OLS假定时，这个方法是可行的。如果一个离散变量只有很少的值，那么我们或许会考虑使用虚拟变量来代表它。另外，可使用回归之外的方法来模型化离散解释变量的作用(Bollen, 1989; Joreskog & Sorbom, 1993)。

一个相关的问题是，隐含的潜在变量从理论上讲是连续的，而研究者观察到的测度却是离散的。使用比例尺度时常出现这种情况(例如，比例从1到7，用于反映隐含的自尊心这一概念)。分析者在使用传统回归模型时，如果这些测度有许多值，同时基本符合OLS回归假定，那一般不会有什么大麻烦。关于克服这些由测度的粗糙性带来的问题，Bollen(1989)和Joreskog以及Sorbom(1993)的著作中有详细的讨论。

第21节 | 调节框架的回顾

调节变量框架提供了在实质层面理解交互作用的有效工具。有些统计学者反对这种方法,因为它从概念上最小化了交互模型中作用的对称性,而且它的关注点离开了那些交互模型中的系数。这些反对声音自有其道理,但用一种理论上合理的概念建立交互框架所带来的收益抵消了这些反对的声音。非调节变量为基础的交互作用概念化强调非加叠性概念,以及分离基于残差化均值的效果(Rosnow & Rosenthal, 1996)。调节方法和后一种方法并不存在哪个更优越的问题,它们是看待数据的不同方法。有些方法论学者(如 Pedhazur, 1997)不仅在统计模型意义上,还根据研究设计中的数据类型及其外在特征来定义交互作用。这些定义是艰涩的,且并不局限于交互分析。在思考乘积项的分析意义时,需要把研究设计的特点也考虑进来。在某些情境下,这些分析可能被误解,反映的是调停而非调节,但对更宽层面的应用而言,本书讨论的框架将被证明是对交互分析进行概念化和操作的有用方法。

注释

[1] 标准分数计算公式为 $Z=\frac{X-\mu}{\sigma}$，X 是原始分数，μ 是总体均值，σ 是总体标准差。——译者注

[2] 当出现抑制变量时，这种对偏相关的理解就具有误导性（见 Cohen & Cohen, 1983）。——译者注

[3] 类型 1 错误即虚无假设为真，但显著性检验拒绝了虚无假设，又称“弃真”。——译者注

[4] 此处原文是 X。根据上下文，可确定是作者或编辑的失误，应当为 Y。——译者注

[5] 在只有主效应的模型中，b_1 和 b_2 是控制其他解释变量时 X 和 Z 对 Y 作用的大小，而控制的意义是指其他解释变量都取各自的均值。在包含乘积项的模型中，b_1 和 b_2 若要直接代表 X 和 Z 对 Y 作用的大小，就必须控制其他解释变量，但此时的控制一般不是其他解释变量取均值，而是指其为 0。

[6] 此处原文为 $6.0+1.0\times 2$，应是编排的失误。——译者注

[7] 原文是“substantial”，据文意应当是 substantially different，即变换前，XZ 和 X 的相关性会很高，而变换后，XZ 和 X 的相关性为 0，两者差异巨大。——译者注

[8] 原文为 European American，直译为美国的欧洲人。出于汉语习惯，译成美国白人，下同。——译者注

[9] 此句原文是“The fact that they are not suggest that there may be three-way interaction”，即指不存在三向交互作用，而事实上，统计结果和图表都表明存在三向交互作用，因此改正。——译者注

[10] 原文如此，疑有误。根据虚拟变量定义，若 D_{Z1}、D_{Z2} 属于同一变量 Z，则可同为 0（代表参照组），但不能同为 1。比如 Z 代表种族（三类），D_{Z1} 代表白人，D_{Z2} 代表黑人（D_{Z1}、D_{Z2} 同时为 0 代表参照组拉美裔），那么 D_{Z1} 和 D_{Z2} 不同时为 1（既是白人又是黑人）。只能计算 $D_{Z1}=1$ 且 $D_{Z2}=0$ 时，或 $D_{Z1}=0$ 且 $D_{Z2}=1$ 时的 X 系数。除非 D_{Z1} 和 D_{Z2} 代表两个独立的虚拟变量，有不同的参照组，比如 D_{Z1} 代表男性（参照组为女性），D_{Z2} 代表白人（参照组为非白人），只有这种情况下才可计算 $D_{Z1}=1$ 且 $D_{Z2}=1$ 时（即男性白人），X 作用大小的系数。——译者注

[11] 原文是 $\beta_1+\beta_2 Q+\beta_5 Z+\beta_7 QZ$，根据文意，明显应当是 $\beta_4 Q$ 而非 $\beta_2 Q$，故改正。——译者注

[12] 标准化系数与非标准化系数的换算公式为 $b'_j = b_j(S_j/S_y)$，$b_j = b'_j(S_y/S_j)$。因此，$0.5 \times (15000/3.0) = 2500$，$0.5 \times (6000/3.0) = 1000$。此例计算基于对黑人、白人分开做收入对教育回归的情况，因此有分种族的收入和教育标准差。——译者注

[13] 假定如下：(1)总体中的残差是正态分布的；(2)潜类连续变量是多变量正态分布的（但潜类乘积项不是）；(3)一个给定的观察分数是真实分数的函数，误差分数符合经典检验理论。

[14] 即残差与自变量相关，导致内生性问题。——译者注

参考文献

Aiken, L., and WEST, S. (1991) *Multiple Regression*. Newbury Park, CA: Sage.

Anderson, N. H. (1981) *Methods of Information Integration Theory*. New York: Academic Press.

Arbuckle, J. L., and Wothke, W. (1999) *AMOS 4.0 User's Guide*. Chicago: Smallwaters.

Berry, W. D., and Feldman, S. (1985) *Multiple Regression in Practice*. Beverly Hills, CA: Sage.

Bishop, Y. M., Feinberg, G., and Holland, P. (1975) *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, MA: MIT Press.

Bohrnstedt, G. W., and Carter, T. M. (1971) "Robustness in regression analysis," in H. L. Costner(ed.) *Sociological Methodology*. San Francisco: Jossey-Bass.

Bohrnstedt, G. W., and Marwell, G. (1978) "The reliability of products of two random variables," in K. F. Schuessler(ed.) *Sociological Methodology*. San Francisco: Jossey-Bass.

Bollen, K. A. (1989) *Structural Equations with Latent Variables*. New York: Wiley.

Bollen, K. A. (1996) "An alternative two stage least squares(2SLS) estimator for latent variable equations." *Psychometrika*, 61:109—121.

Bollen, K. A., and Paxton, P. (1998) "Interaction of latent variables in structural equation models," *Structural Equation Modeling* 5: 267—293.

Bryk, A., and Raudenbush, S. (2002) *Hierarchical Linear Models: Applications and Data Analysis Methods*. Thousand Oaks, CA: Sage.

Busemeyer, J. R., and Jones, L. (1983) "Analysis of multiplicative combination rules when the cause variables are measured with error." *Psychological Bulletin* 93:549—562.

Cohen, J., and Cohen, P. (1975) *Applied Multiple Regression for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum.

Cohen, J., and Cohen, P. (1983) *Applied Multiple Regression for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum.

Cronbach, L. (1987) "Statistical tests for moderator variables: Flaws in

analysis recently proposed." *Psychological Bulletin* 102:414—417.

Cronbach, L., and Gleser, G. C. (1957) *Psychological Tests and Personal Decisions* (2nd ed.). Urbana: University of Illinois Press.

Cronbach, L. J., and Snow, R. E. (1981) *Aptitudes and Instructional Methods: A Handbook for Research on Interactions*. New York: Irvington.

Duncan, T. E., Duncan, S. E., Strycker, L. A., Li, F., and Alpert, A. (1999) *An Introduction to Latent Variable Growth Curve Modeling: Concepts, Issues, and Applications*. Mahwah, NJ: Erlbaum.

Evans, M. T. (1985) "A Monte Carlo study of the effects of correlated variance in moderated multiple regression analysis." *Organizational Behavior and Human Decision Process* 36:305—323.

Friedrich, R. (1982) "In defense of multiplicative terms of multiple regression equations." *American Journal of Political Sciences* 26:797—833.

Fuller, W. A., and Hidiroglu, M. A. (1978) "Regression estimation after correction for attenuation." *Journal of American Statistical Association* 73:99—104.

Ganzach, Y. (1997) "Misleading interaction and curvilinear terms." *Psychological Methods* 2:235—247.

Greene, W. H. (1997) *Econometric Analysis*. Upper Saddle River, NJ: Prentice Hall.

Hamilton, L. C. (1992) *Regression with Graphics: A Second Course in Applied Statistics*. Belmont, CA: Brooks-Cole.

Hardy, K. (1993) *Regression with Dummy Variables*. Newbury Park, CA: Sage.

Heise, D. R. (1986) "Estimating non-linear models: Correcting for measurement error." *Sociological Methods of Research* 14:447—472.

Holland, B. S., and Copenhaver, M. (1988) "Improved Bonferroni-type multiple testing procedures." *Psychological Bulletin* 104:145—149.

Holm, S. (1979) "A simple sequentially rejective multiple test procedure." *Scandinavian Journal of Statistics* 6:65—70.

Hosmer, D.W., and Lemeshow, S. (1989) *Applied Logistic Regression*. New York: Wiley.

Jaccard, J. (1998) *Interaction Effects in Factorial Analysis of Variance*. Thousand Oaks, CA: Sage.

Jaccard, J., and Dodge. T. (2003) "Specification of contingent effects in lin-

ear models," in M. Hardy and A. Bryman(eds.), *Handbook of Data Analysis*. Thousand Oaks, CA: Sage, in press.

Jaccard, J., Turrisi, R., and Wan, C. (1990) *Interaction Effects in Multiple Regression*. Newbury Park, CA: Sage.

Jaccard, J., and Wan, C. (1996) *LISREL Approaches to Interaction Effects in Multiple Regression*. Thousand Oaks, CA: Sage.

James, L., and Tetrick, L. (1984) "A multivariate test of homogeneity of regression weights for correlated data." *Educational and Psychological Measurement* 44:769—780.

Johnson, P. O., and Neyman, J. (1936) "Tests of certain linear hypothesis and their application to some educational problems." *Statistical Research Memoirs* 1:57—93.

Joreskog, K., and Sorbom, D. (1993) *LISREL Ⅷ*. Chicago: Scientific Software.

Judd, C. M., Kenny, D. A., and McClelland, G. (2002) "Estimating and testing mediation and moderation in within subject designs." *Psychological Methods* 6:115—134.

Kleinbaum, D. G. (1992) *Logistic Regression: A Self Learning Text*. New York: Springer.

Kline, R. B. (1998) *Principles and Practice of structural Equation Modeling*. New York: Guilford Press.

Lewis-Beck, M. S. (1980) *Applied Regression: An Introduction*. Beverly Hilly, CA: Sage.

Lubinski, D., and Humphreys, L. (1990) "Assessing spurious 'moderator effects': Illustrated substantively with the hypothesized('synergistic') relation between spatial and mathematical ability." *Psychological Bulletin* 107:385—393.

Maxwell, S., and Delaney, H. (1993) "Bivariate median splits and spurious statistical significance." *Psychological Bulletin* 113:181—190.

McClelland, G. H., and Judd, C. M. (1993) "Statistical difficulties of detecting interactions and moderator effects." *Psychological Bulletin* 114: 376—390.

Pedhazur, E. (1997) *Multiple Regression in Behavioral Research: Explanation and Prediction*. New York: Harcourt Brace.

Potthoff, R. F. (1964) "On the Johnson-Neyman technique and some extensions thereof." *Psychometrika* 29:241—256.

Rogers, W. M. (2002) "Theoretical and mathematical constraints of interactive regression models." *Organizational Research Methods* 5: 212—230.

Rosnow, R. L., and Rosenthal, R. (1996) "Contrasts and interaction effects redux: Five easy pieces." *Psychological Sciences* 7:253—257.

Schroeder, L. D., Sjoquist, D. L., and Stephan, P. (1986) *Understanding Regression Analysis: An Introductory Guide*. Beverly Hills, CA: Sage.

Seaman, M. A., Levin, K. R., and Serlin, R. C. (1991) "New developments in pairwise multiple comparisons: Some powerful and practicable procedures." *Psychological Bulletin* 110:577—586.

Vandenberg, R. J., and Lance, C. (2000) "A review and synthesis of the measurement in variance literature: Suggestions, practices, and recommendations for organizational research." *Organizational Research Methods* 3:4—69.

Wall, M., and Amemiya, Y. (2000) "Estimation for polynomial structural equation models." *Journal of American Statistical Association* 95: 929—940.

Wegnar, B. (1982) *Social Attitudes and Psychological Measurement*. Hillsdale, NJ: Lawrence Erlbaum.

Westfall, P. H., Tobias, R., Rom, D., Wolfinger, R., and Hochberg, R. D. (1999) *Multiple Comparisons and Multiple Tests Using the SAS System*. Cary, NC: SAS Institute.

Wilcox, R. R. (1997) *Introduction to Robust Estimation and Hypothesis Testing*. New York: Academic Press.

Wilcox, R. R. (2001) *Fundamentals of Modern Statistical Methods*. New York: Springer-Verlag.

Wilkinson, L. (1999) "Statistical methods in psychology journals: Guidelines and explanations." *American Psychologist* 54:594—604.

译名对照表

backward elimination test	逆向消除法
bidirectional or reciprocal causal relationship	互为因果关系
bilinear interaction	双线性交互作用
mean centering	对中
clustered	聚类
dynamics	机制
experimentwise error	实验误差
factorial design	析因设计
hierarchical regression	层级回归
hypothesis	假设
interaction effects	交互作用
level	类别
metric	度量
moderated relationship	调节关系
omnibus interaction test	整体交互作用检验
overfit	过度拟合
product term	乘积项
robust	稳健
scale	尺度
squared semipartial correlation	半偏相关系数平方
simple main effects	简单主效应
squared multiple correlation	复相关系数平方(R^2)
two-way interaction	双向交互作用
three-way interaction	三向交互作用
three-term equation	三项目方程
variance heterogeneity	方差异质性



格致方法·定量研究系列

微信



微博

上架建议：社会研究方法


ISBN 978-7-5432-2597-8

9 787543 225978 >

定价：25.00元

易文网：www.ewen.co

格致网：www.hibooks.cn